Benjamin A. **Pierce**

Genética

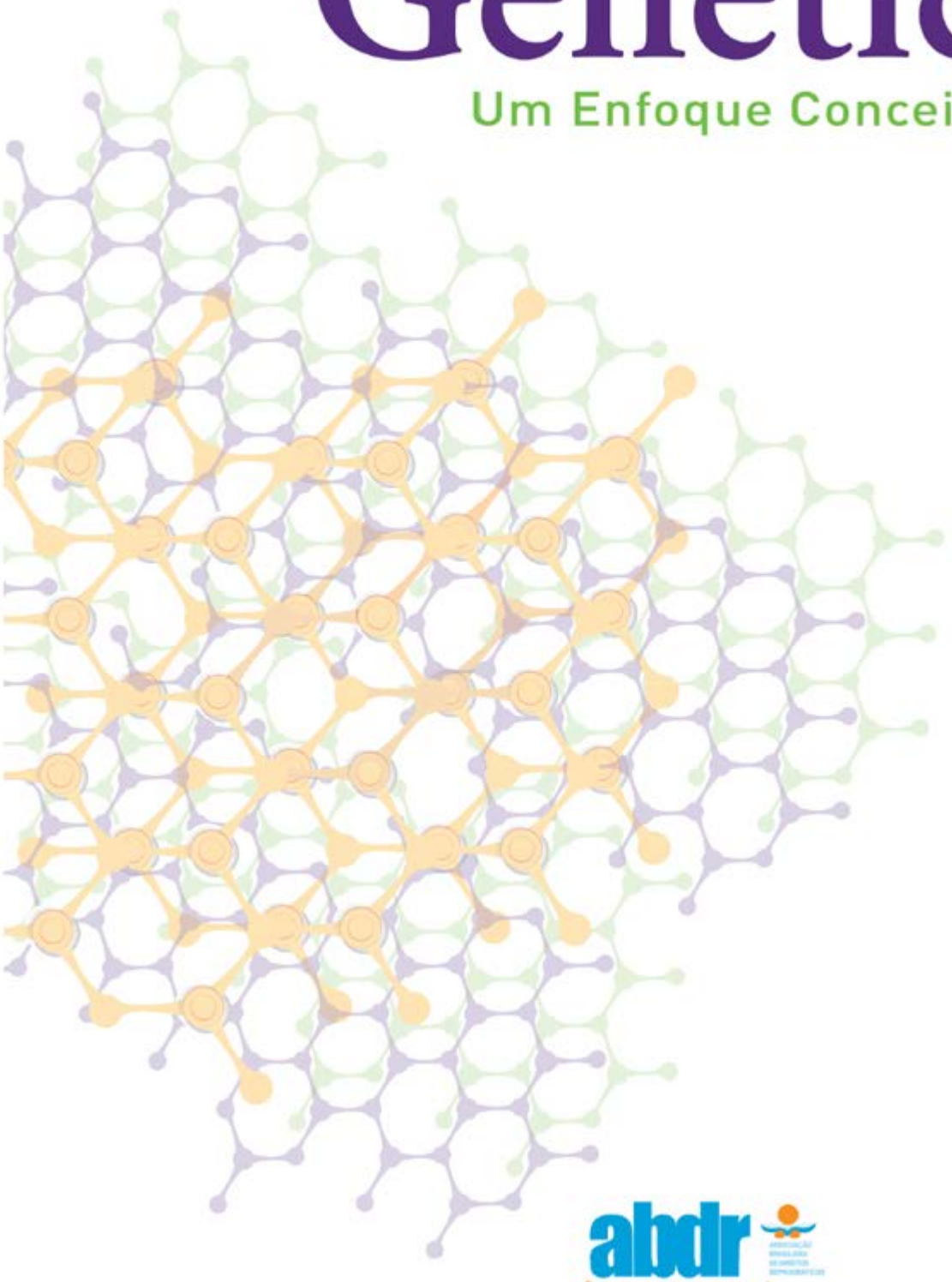
Um Enfoque Conceitual

Quinta edição

gen | GUANABARA KOOGAN

Genética

Um Enfoque Conceitual



abdr
Respeite o direito autoral

O GEN | Grupo Editorial Nacional, a maior plataforma editorial no segmento CTP (científico, técnico e profissional), publica nas áreas de saúde, ciências exatas, jurídicas, sociais aplicadas, humanas e de concursos, além de prover serviços direcionados a educação, capacitação médica continuada e preparação para concursos. Conheça nosso catálogo, composto por mais de cinco mil obras e três mil e-books, em www.grupogen.com.br.

As editoras que integram o GEN, respeitadas no mercado editorial, construíram catálogos inigualáveis, com obras decisivas na formação acadêmica e no aperfeiçoamento de várias gerações de profissionais e de estudantes de Administração, Direito, Engenharia, Enfermagem, Fisioterapia, Medicina, Odontologia, Educação Física e muitas outras ciências, tendo se tornado sinônimo de seriedade e respeito.

Nossa missão é prover o melhor conteúdo científico e distribuí-lo de maneira flexível e conveniente, a preços justos, gerando benefícios e servindo a autores, docentes, livreiros, funcionários, colaboradores e acionistas.

Nosso comportamento ético incondicional e nossa responsabilidade social e ambiental são reforçados pela natureza educacional de nossa atividade, sem comprometer o crescimento contínuo e a rentabilidade do grupo.

Genética

Um Enfoque Conceitual

Benjamin A. Pierce

Southwestern University

Tradução
Beatriz Araujo do Rosário

Revisão Técnica
Claudia Vitoria de Moura Gallo
Professora-associada do Departamento de Genética da Universidade do Estado do Rio de Janeiro (Uerj).
Doutora pela Université Paris VII.

Quinta edição

gen | GUANABARA KOOGAN

■ O autor deste livro e a EDITORA GUANABARA KOOGAN LTDA. empenharam seus melhores esforços para assegurar que as informações e os procedimentos apresentados no texto estejam em acordo com os padrões aceitos à época da publicação, *e todos os dados foram atualizados pelo autor até a data da entrega dos originais à editora*. Entretanto, tendo em conta a evolução das ciências da saúde, as mudanças regulamentares governamentais e o constante fluxo de novas informações sobre terapêutica medicamentosa e reações adversas a fármacos, recomendamos enfaticamente que os leitores consultem sempre outras fontes fidedignas, de modo a se certificarem de que as informações contidas neste livro estão corretas e de que não houve alterações nas dosagens recomendadas ou na legislação regulamentadora. *Adicionalmente, os leitores podem buscar por possíveis atualizações da obra em http://gen-io.grupogen.com.br*.

■ O autor e a editora se empenharam para citar adequadamente e dar o devido crédito a todos os detentores de direitos autorais de qualquer material utilizado neste livro, dispondo-se a possíveis acertos posteriores caso, inadvertida e involuntariamente, a identificação de algum deles tenha sido omitida.

■ Traduzido de:
GENETICS: A CONCEPTUAL APPROACH, FIFTH EDITION
First published in the United States by
W.H. FREEMAN AND COMPANY, New York
Copyright © 2014, 2012, 2008, 2006 by W. H. Freeman and Company.
All Rights Reserved.

Publicado originalmente nos Estados Unidos por
W.H. FREEMAN AND COMPANY, New York
Copyright © 2014, 2012, 2008, 2006 by W. H. Freeman and Company.
Todos os Direitos Reservados.

ISBN: 978-1-4641-0946-1

■ Direitos exclusivos para a língua portuguesa
Copyright © 2016 by
EDITORA GUANABARA KOOGAN LTDA.
Uma editora integrante do GEN | Grupo Editorial Nacional
Travessa do Ouvidor, 11
Rio de Janeiro – RJ – CEP 20040-040
Tel.: (21) 3543-0770 | Fax: (21) 3543-0896
www.editoraguanabara.com.br | www.grupogen.com.br | editorial.saude@grupogen.com.br

■ Reservados todos os direitos. É proibida a duplicação ou reprodução deste volume, no todo ou em parte, em quaisquer formas ou por quaisquer meios (eletrônico, mecânico, gravação, fotocópia, distribuição pela Internet ou outros), sem permissão, por escrito, da EDITORA GUANABARA KOOGAN LTDA.

■ Capa: Editorial Saúde
Produção digital: Geethik

■ Ficha catalográfica

P674g
5. ed.

Pierce, Benjamin A.
Genética: um enfoque conceitual / Benjamin A. Pierce; tradução Beatriz Araujo do Rosário. - 5. ed. - Rio de Janeiro: Guanabara Koogan, 2016.
il.

Tradução de: Genetics: a conceptual approach
ISBN 978-85-277-2932-1

1. Genética. I. Rosário, Beatriz Araujo do.

16-29599

CDD: 575.1
CDU: 575

Dedicatória

A meus pais, Rush e Amanda Pierce;
a meus filhos, Sarah Pierce Dumas e Michael Pierce;
e à minha parceira genética, amiga e alma gêmea
por 33 anos, Marlene Tyrrell.

Carta do Autor

Ainda me lembro do entusiasmo que senti ao fazer meu primeiro curso de genética. À época, estava intrigado com os princípios da hereditariedade, que possibilitavam prever como seriam os descendentes antes mesmo de nascerem. Além de me fascinar o fato de aprender que esses conceitos se fundamentavam na química de uma elegante molécula chamada DNA, eu buscava descobrir que essa ciência é a base para o processo de evolução, responsável pela infinita diversidade e beleza da vida. Tais elementos da genética me impressionam até hoje, e a possibilidade de transmitir aos estudantes esse entusiasmo é uma das melhores virtudes em ensinar essa disciplina.

Esse foi justamente o objetivo original que me levou a escrever – após tantos anos em sala de aula – a obra *Genética | Um Enfoque Conceitual*. Desde a primeira à quinta edição, ela busca motivar os leitores ao estudo dessa disciplina, tendo como base os conceitos e a solução de problemas relacionados ao tema.

Nesta edição, compartilho o que tenho aprendido em mais de três décadas de ensino de genética; além disso, oriento os estudantes nas áreas em que eles costumam ter dificuldades e trago à tona fatos – tanto passados quanto presentes – relacionados a pessoas, locais e experimentos de genética, tudo para manter esse assunto sempre relevante, interessante e vivo.

Meu objetivo é ajudar os leitores a se ater aos detalhes necessários e aprender os conceitos e as habilidades para solucionar problemas, enquanto os incentivo a enxergar a elegância e beleza dessa disciplina.

Minha sala está sempre aberta na Southwestern University, onde meus alunos constantemente me visitam para compartilhar

as próprias abordagens sobre o aprendizado, os artigos que leram sobre genética, suas experiências, seus conceitos e conquistas. Aprendo com esses estudantes, tal como eles aprendem comigo, e gostaria muito de aprender também com você, seja por *e-mail* (pierceb@southwestern.edu), telefone (512-863-1974) ou pessoalmente (Southwestern University, Georgetown, Texas).

Ben Pierce

Professor Universitário de Biologia

Ocupante da Lillian Nelson Pratt Chair

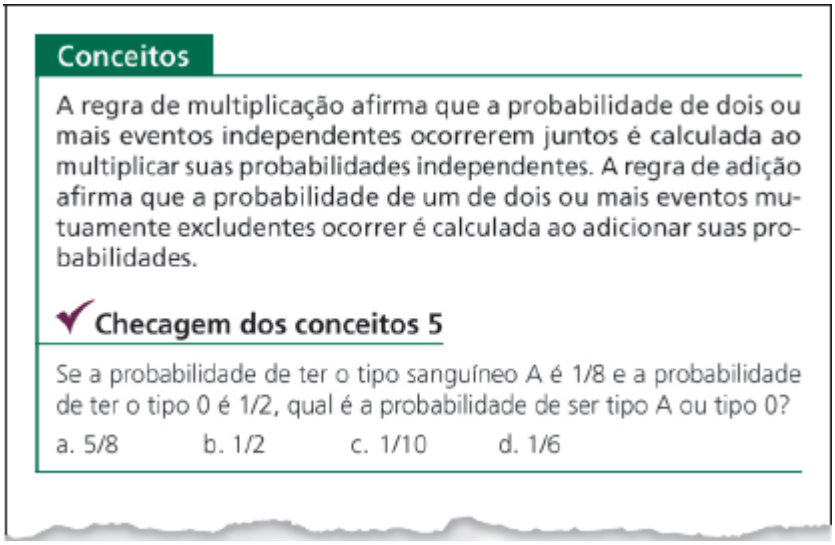
Southwestern University

Prefácio

O principal objetivo de *Genética | Um Enfoque Conceitual* sempre foi ajudar os estudantes a desvendar os principais conceitos da genética e a estabelecer conexões entre eles. Nas edições anteriores, o estilo de escrita acessível, as ilustrações didáticas e as características pedagógicas ajudaram os estudantes a desenvolver uma compreensão mais completa da genética, recursos que foram mantidos nesta quinta edição.

Principais características

- **Conceitos-chave e conexões.** Em todo o livro, recursos pedagógicos ajudam os estudantes a ter como foco os principais conceitos de cada tópico.
  - Os boxes ***Conceitos*** resumem os pontos principais de cada capítulo. As perguntas dos itens *Checagem dos conceitos* apresentam-se em formato de múltipla escolha e discursivas e estão listadas no final de cada capítulo com o objetivo de ajudar os estudantes a avaliar rapidamente sua compreensão do material que acabaram de ler.
  - Os *Conceitos conectantes* comparam os processos e mostram no que eles se diferem, além de integrarem ideias-chave por seções e capítulos, para ajudar os estudantes a compreender como os diferentes tópicos de genética se relacionam. Todos os principais conceitos estão elencados no *Resumo dos conceitos*, ao final de cada capítulo.



Conceitos

A regra de multiplicação afirma que a probabilidade de dois ou mais eventos independentes ocorrerem juntos é calculada ao multiplicar suas probabilidades independentes. A regra de adição afirma que a probabilidade de um de dois ou mais eventos mutuamente excludentes ocorrer é calculada ao adicionar suas probabilidades.

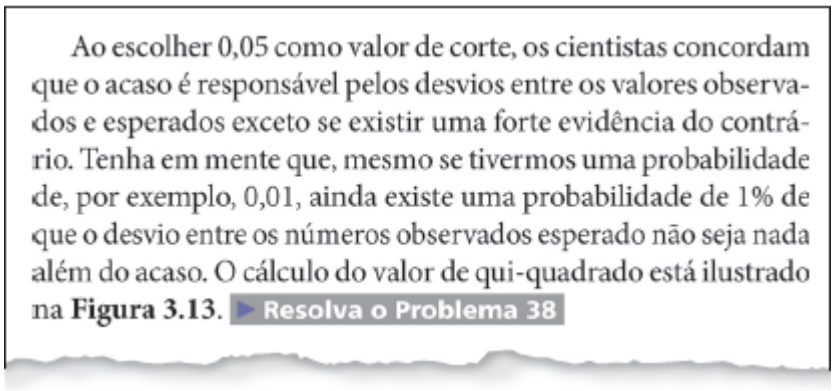
Checagem dos conceitos 5

Se a probabilidade de ter o tipo sanguíneo A é 1/8 e a probabilidade de ter o tipo 0 é 1/2, qual é a probabilidade de ser tipo A ou tipo 0?

a. 5/8 b. 1/2 c. 1/10 d. 1/6

- **Acessibilidade.** O estilo de escrita simplificado tem sido uma das características favoritas deste livro para estudantes e professores. Além de conduzir cuidadosamente os alunos pelos conceitos principais da genética, inicio cada capítulo com uma história introdutória sobre o assunto a ser estudado. Mais de um terço dessas histórias são novas e, em sua totalidade, destacam exemplos relevantes de doenças ou outros fenômenos biológicos, servindo como uma prévia do que os estudantes aprenderão nos capítulos.
- **Programa de ilustração claro e simples.** As figuras atrativas e instrutivas provaram ser uma ferramenta eficiente para o aprendizado dos estudantes durante as edições anteriores e continuam sendo um destaque nesta nova edição. Cada ilustração foi cuidadosamente desenhada para ressaltar os pontos principais e conduzir, passo a passo, os leitores pelos experimentos e processos. A maioria das figuras contém textos que os norteiam nas apresentações gráficas, e as ilustrações de experimentos reforçam o método científico ao proporem primeiro uma hipótese, destacarem posteriormente os métodos e resultados e apresentarem conclusões que corroboram os conceitos explicados no texto
- **Ênfase na solução de problemas.** Um dos fatos que observei em mais de três décadas de magistério é que os estudantes

aprendem melhor genética ao solucionarem problemas. Por meio de exemplos, equações ou experimentos, eles conseguem ver em ação os conceitos aprendidos e reforçar as ideias explicadas. No livro, ajudo os alunos a desenvolver as habilidades para solucionar de várias maneiras esses problemas. Os *Problemas desenvolvidos* foram recém-reformulados (ver *Conteúdo novo e reorganizado*, na página seguinte) para conduzir os estudantes em cada etapa de um conceito difícil. O ícone *Resolva o Problema* destaca sempre aqueles do final do capítulo em que os estudantes podem testar sua compreensão a respeito do que acabaram de ler; todos esses problemas contêm respostas no final da obra. Além disso, uma ampla variedade de problemas encerra cada capítulo, sendo organizados por seção e classificados em *Questões de compreensão*, *Questões e problemas aplicados* e *Questões desafiadoras*. Algumas delas foram elaboradas com base em dados reais da literatura científica e estão marcadas com o ícone *Análise de dados*.



Ao escolher 0,05 como valor de corte, os cientistas concordam que o acaso é responsável pelos desvios entre os valores observados e esperados exceto se existir uma forte evidência do contrário. Tenha em mente que, mesmo se tivermos uma probabilidade de, por exemplo, 0,01, ainda existe uma probabilidade de 1% de que o desvio entre os números observados esperado não seja nada além do acaso. O cálculo do valor de qui-quadrado está ilustrado na **Figura 3.13**. ▶ Resolva o Problema 38

Novidades da quinta edição

Conteúdo novo e reorganizado

Partindo do princípio de que nossa compreensão está em constante mudança no que diz respeito à herança, natureza molecular das informações genéticas, epigenética e evolução genética, esta quinta edição discute as descobertas recentes sobre esses temas.

Visando integrar e expandir o conteúdo disposto em cinco capítulos da edição anterior, a obra conta com um novo capítulo sobre epigenética (Capítulo 21). No entanto, discussões mais concisas sobre o tema também são pontualmente encontradas nos Capítulos 5, 11, 17, 22 e 23.

O Capítulo 11 trata da *Estrutura do Cromossomo e DNA de Organelas* (tema que, na edição anterior, integrava o Capítulo 21).

A ordem dos Capítulos 8 e 9 da edição anterior (*Sistemas Genéticos Bacterianos e Virais* e *Variação Cromossômica*) foi trocada para que a *Variação Cromossômica* agora siga o Capítulo 7 (*Ligação Gênica, Recombinação e Mapeamento de Genes Eucarióticos*).

Outros conteúdos novos e atualizados incluem:

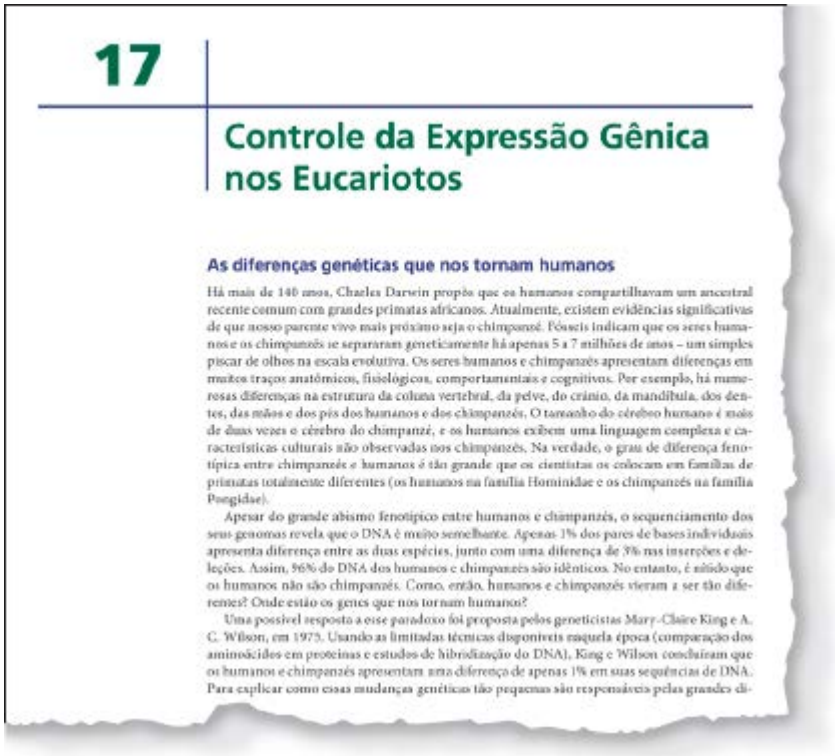
- Debates sobre *A natureza molecular dos alelos* e *A probabilidade condicional* (Capítulo 3)
- Ampla revisão da *Determinação do sexo na Drosophila melanogaster* e da compensação da dose (Capítulo 4)
- Discussão ampliada sobre interferência genética (Capítulo 7)
- Seção significativamente revisada sobre rearranjos cromossômicos (Capítulo 8)
- Discussão ampliada sobre a descoberta da estrutura de DNA, incluindo a contribuição de Franklin (Capítulo 10)
- Debates sobre a origem evolutiva dos íntrons, os CRISPR RNAs, os RNAs longos não codificados e a regulação gênica (Capítulo 14)
- Discussão sobre o *Decaimento no-go* (Capítulo 15)
- Debate sobre *Acentuadores bacterianos* (Capítulo 16)
- Estudo atualizado e ampliado sobre as mudanças na estrutura da cromatina, bem como abordagem sobre a *Imunoprecipitação da cromatina* (Capítulo 17)
- Discussão sobre o *Sequenciamento Illumina* e revisão sobre o *Fingerprinting do DNA* (Capítulo 19)
- Atualização e ampliação dos *Estudos de associação genômica ampla* (Capítulo 20) e da *Metagenômica* (Capítulo 20)
- Estudo comparativo sobre o desenvolvimento da *Drosophila* e das flores (Capítulo 22)
- Debate sobre *Evolução por meio de mudanças na regulação gênica* (Capítulo 26).

Os *Problemas desenvolvidos* foram reformulados para ajudar os estudantes em conceitos quantitativos de difícil compreensão, por meio das *Estratégias para a solução* e das *Etapas para a solução* em cada um deles. Além disso, há balões

contendo dicas e lembretes que levam o leitor a rever e reforçar conceitos importantes estudados no capítulo.

Alguns dos novos problemas ao final dos capítulos contêm figuras – recurso útil para a solução dos problemas, pois, por meio delas, os estudantes têm mais facilidade em compreender os conceitos e processos-chave.

As novas histórias introdutórias, narradas brevemente no início de cada capítulo, ilustram a relevância de determinado conceito genético que os estudantes aprenderão. Essas histórias – uma das características favoritas das edições anteriores – possibilitam que os estudantes vislumbrem o campo da genética atual e ajudam a atrair a atenção deles para os capítulos. Entre os novos tópicos de introdução, citamos: *O mistério das impressões digitais desaparecidas*; *Como criar bananas de melhor qualidade*; *As diferenças genéticas que nos tornam humanos*; e *Como a dieta do seu avô pode afetar a sua saúde*. Os problemas ao final dos capítulos tratam especificamente de conceitos discutidos na maioria das histórias introdutórias, tanto as antigas quanto as novas.



17

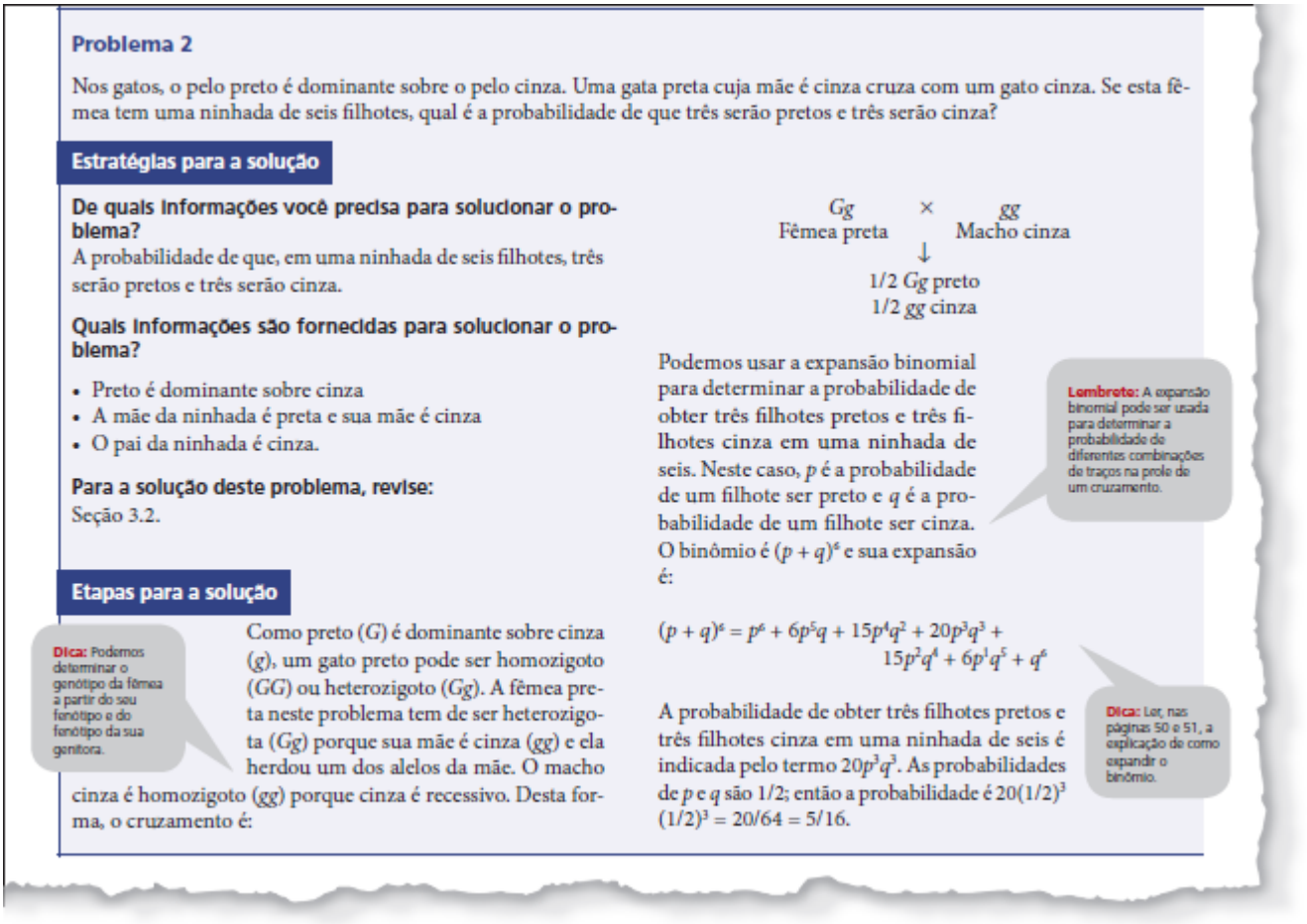
Controle da Expressão Gênica nos Eucariotos

As diferenças genéticas que nos tornam humanos

Há mais de 140 anos, Charles Darwin propôs que os humanos compartilhavam um ancestral recente comum com grandes primatas africanos. Atualmente, existem evidências significativas de que nosso parente vivo mais próximo seja o chimpanzé. Fósseis indicam que os seres humanos e os chimpanzés se separaram geneticamente há apenas 5 a 7 milhões de anos – um simples piscar de olhos na escala evolutiva. Os seres humanos e chimpanzés apresentam diferenças em muitos traços anatômicos, fisiológicos, comportamentais e cognitivos. Por exemplo, há numerosas diferenças na estrutura da coluna vertebral, da pelve, do crânio, da mandíbula, dos dentes, das mãos e dos pés dos humanos e dos chimpanzés. O tamanho do cérebro humano é mais de duas vezes o cérebro do chimpanzé, e os humanos exibem uma linguagem complexa e características culturais não observadas nos chimpanzés. Na verdade, o grau de diferença fenotípica entre chimpanzés e humanos é tão grande que os cientistas os colocam em famílias de primatas totalmente diferentes (os humanos na família Hominidae e os chimpanzés na família Pongidae).

Apesar do grande abismo fenotípico entre humanos e chimpanzés, o sequenciamento dos seus genomas revela que o DNA é muito semelhante. Apenas 1% dos pares de bases individuais apresenta diferença entre as duas espécies, junto com uma diferença de 3% nas inserções e deleções. Assim, 96% do DNA dos humanos e chimpanzés são idênticos. No entanto, é nítido que os humanos não são chimpanzés. Como, então, humanos e chimpanzés vieram a ser tão diferentes? Onde estão os genes que nos tornam humanos?

Uma possível resposta a esse paradoxo foi proposta pelos geneticistas Mary-Claire King e A. C. Wilson, em 1975. Usando as limitadas técnicas disponíveis naquela época (comparação dos aminoácidos em proteínas e estudos de hibridização do DNA), King e Wilson concluíram que os humanos e chimpanzés apresentam uma diferença de apenas 1% em suas sequências de DNA. Para explicar como essas mudanças genéticas tão pequenas são responsáveis pelas grandes di-



Problema 2

Nos gatos, o pelo preto é dominante sobre o pelo cinza. Uma gata preta cuja mãe é cinza cruza com um gato cinza. Se esta fêmea tem uma ninhada de seis filhotes, qual é a probabilidade de que três serão pretos e três serão cinza?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
A probabilidade de que, em uma ninhada de seis filhotes, três serão pretos e três serão cinza.

Quais informações são fornecidas para solucionar o problema?

- Preto é dominante sobre cinza
- A mãe da ninhada é preta e sua mãe é cinza
- O pai da ninhada é cinza.

Para a solução deste problema, revise:
Seção 3.2.

Etapas para a solução

Dica: Podemos determinar o genótipo da fêmea a partir do seu fenótipo e do fenótipo da sua genitora.

Como preto (*G*) é dominante sobre cinza (*g*), um gato preto pode ser homozigoto (*GG*) ou heterozigoto (*Gg*). A fêmea preta neste problema tem de ser heterozigota (*Gg*) porque sua mãe é cinza (*gg*) e ela herdou um dos alelos da mãe. O macho cinza é homozigoto (*gg*) porque cinza é recessivo. Desta forma, o cruzamento é:

$$\underset{\text{Fêmea preta}}{Gg} \times \underset{\text{Macho cinza}}{gg}$$
$$\downarrow$$
$$1/2\ Gg \text{ preto}$$
$$1/2\ gg \text{ cinza}$$

Podemos usar a expansão binomial para determinar a probabilidade de obter três filhotes pretos e três filhotes cinza em uma ninhada de seis. Neste caso, p é a probabilidade de um filhote ser preto e q é a probabilidade de um filhote ser cinza. O binômio é $(p + q)^6$ e sua expansão é:

Lembrete: A expansão binomial pode ser usada para determinar a probabilidade de diferentes combinações de traços na prole de um cruzamento.

$$(p + q)^6 = p^6 + 6p^5q + 15p^4q^2 + 20p^3q^3 + 15p^2q^4 + 6p^1q^5 + q^6$$

A probabilidade de obter três filhotes pretos e três filhotes cinza em uma ninhada de seis é indicada pelo termo $20p^3q^3$. As probabilidades de p e q são 1/2; então a probabilidade é $20(1/2)^3(1/2)^3 = 20/64 = 5/16$.

Dica: Ler, nas páginas 50 e 51, a explicação de como expandir o binômio.

Agradecimentos

Sou grato a várias pessoas que me ajudaram tanto nesta edição de *Genética | Um Enfoque Conceitual* quanto nas anteriores. Aprendi muito com meus professores Ray Canham, que me apresentou à genética pela primeira vez e incutiu em mim esse amor para toda a vida, e Jeff Mitton, que me ensinou a arte da pesquisa genética. Aprendi também com muitos alunos de genética que frequentaram minhas aulas nas últimas três décadas, primeiro na Connecticut College, depois na Baylor University e agora na Southwestern University. A inteligência, o entusiasmo, a curiosidade e o humor dessas pessoas são a fonte de motivação e alegria durante toda a minha vida como professor.

Também aprendi com estudantes de todo o mundo, que usaram as edições anteriores e, por *e-mails* ou telefonemas, gentilmente compartilharam comigo suas opiniões sobre o livro e sugestões de melhoria.

Sou grato pelos incríveis colegas que me acompanham diariamente na Southwestern University; a amizade, a orientação e o bom humor de cada um deles são o apoio para meu trabalho. As pequenas turmas, a interação com os estudantes e a capacidade de integração entre ensino e pesquisa tornaram o dia a dia nessa universidade gratificante do ponto de vista pessoal e profissional. Agradeço a James Hunt, reitor da Southwestern University e decano da Brown College, por me ajudar a criar esse ambiente acadêmico de apoio, pela amizade e pelo coleguismo.

Escrever um livro moderno de ciência requer esforço em grupo – e, a esse respeito, fui agraciado com uma incrível equipe na W. H. Freeman and Company. A editora da Life Sciences, Susan Winslow, é uma defensora do livro por muitos anos; admiro sua criatividade, seu discernimento e apoio. Lauren Schultz, editora de aquisições sênior, foi para a equipe da Life Sciences uma incrível orientadora, dando ideias, incentivo e suporte durante todo este trabalho. Foi uma satisfação trabalhar com a editora de desenvolvimento Anna Bristow, minha parceira diária para moldar esta edição de *Genética*. Anna é uma excelente editora, que, com uma incrível capacidade de organizar e atenta aos detalhes, manteve o projeto nos trilhos e contribuiu de diversas maneiras para esta edição.

A editora sênior de projeto Georgia Lee Hadler, da W. H. Freeman, gerenciou com destreza a produção desta edição, assim como das anteriores. Sua dedicação para a excelência em todas as fases do processo foi fundamental para tornar este livro um sucesso. Jeanine Furino trabalhou como copidesque da obra em inglês e contribuiu com valiosas sugestões editoriais. Sou grato a Dragonfly Media Group, por criar e revisar as ilustrações do livro, e Matthew McAdams, por coordenar o programa de ilustração. Agradeço também a Paul Rohloff, da W. H. Freeman, e Assunta Petrone, da codeMantra, pela coordenação das fases de composição e produção, e a Diana Blume, que desenvolveu o projeto do livro e a capa para esta edição em inglês.

Agradeço a Christine Buese e Jacqui Wong a pesquisa das fotos; a Allison Michael, Elaine Palucki, Alexi Garrett, Adam

Feil e Chris Efstratiou, o desenvolvimento da mídia e dos suplementos que acompanham o livro em inglês. Sou grato a Jung Choi e Mark McCallum por escreverem as soluções para os novos problemas do final dos capítulos. Agradeço também a Joseph Ahlander, Ellen France, Robert Fowler, Brian Kreiser, Joshua Loomis, Amy McMillan, Marcie Moehnke, Douglas Thrower e Daniel Williams, por desenvolverem e revisarem as questões de avaliação.

Aos representantes de vendas, gerentes regionais e especialistas regionais de vendas da W. H. Freeman, sou grato por terem divulgado meu livro para os instrutores de genética em todo o mundo. Tive grande prazer de trabalhar com essa equipe de vendas; a opinião, o trabalho árduo e o bom serviço de cada um deles são responsáveis pelo sucesso dos livros da Freeman.

Agradeço a assistência de vários colegas, que revisaram este livro, gentilmente cedendo seu conhecimento técnico e sua experiência de ensino. Qualquer falha eventual é de minha inteira responsabilidade.

Marlene Tyrrell – minha esposa e melhor amiga há 33 anos – nossa filha, Sarah, seu marido, Matt, nosso filho, Michael, e sua esposa, Amber, são a fonte de amor, apoio e inspiração para tudo o que faço.

Minha gratidão aos revisores desta quinta edição de *Genética | Um Enfoque Conceitual*.

Amy Abdulovic-Cui
Augusta State University

Joseph Ahlander
Northeastern State University

David Aiello
Austin College

Preston Aldrich
Benedictine University

Kirk Anders
Gonzaga University

Jeanne Andreoli
Marygrove College

Anthony Arment
Central State University

Minoo Askari
Pellissippi State Community College

Andrea Bailey
Brookhaven Community College

Paul W. Bates
University of Minnesota Duluth

Keith Barlow
Guilford Technical Community College

Phillip Barnes
Connecticut College

Christine Beatty
Benedictine University

John Belote
Syracuse University

Spencer Benson
University of Maryland

Edward Berger
Dartmouth College

Laura Bermingham
University of Vermont

Aimee Bernard

University of Colorado Denver

Indrani Bose
Western Carolina University

John Braverman
Saint Joseph's University

David Buchanan
North Dakota State University

Gerald L. Buldak
Loyola University Chicago

Alyssa C. Bumbaugh
Penn State University

Dan Caprioglio
Colorado State University

Maria V. Cattell
University of Colorado

Richard Duhrkopf
Baylor University

Joel Chandlee
University of Rhode Island

Henry Chang
Purdue University

Cynthia Church
Metropolitan State University of Denver

Sarah Crawford
Southern Connecticut State University

Marilyn Cruz-Alvarez
Florida Gulf Coast University

Sandra Davis
University of Indianapolis

Sarwan Dhir
Fort Valley State University

David Donnell
The Citadel

Chuanguang Du
Montclair State University

Cheryld L. Emmons
Alfred University

William Ettinger
Gonzaga University

Sarah Evans
Friends University

Victor Fet
Marshall University

Ted Fickel
American Jewish University

Robert G. Fowler

San Jose State University

Thomas Fowler
Southern Illinois University Edwardsville

Edison Fowlks
Hampton University

Dennis Frisby
Cameron University

Laura Frost
Point Park University

J. Yvette Gardner
Clayton State University

William Gilliland
DePaul University

Elliott S. Goldstein
Arizona State University

Steven W. Gorsich
Central Michigan University

Anjali Gray
Lourdes University

Bradley Hersh
Allegheny College

Debra Hinson
Dallas Baptist University

Peter Hoffman
Notre Dame of Maryland University

Margaret Hollingsworth
University at Buffalo

Carina Endres Howell
Lock Haven University

Li Huang
Montana State University

Mary Huff
Bellarmine University

Colin Hughes
Florida Atlantic University

Jeffrey Hughes
Millikin University

Cristina Iftode
Rowan University

Jeba Inbarasu
Metropolitan Community College, South Omaha Campus

Diana Ivankovic
Anderson University

Aaron Johnson
University of Colorado School of Medicine

David H. Kass

Eastern Michigan University

Todd Kelson
Brigham Young University–Idaho

Cathy Silver Key
North Carolina Central University

Christopher Korey
College of Charleston

Margaret J. Kovach
University of Tennessee at Chattanooga

Brian Kreiser
University of Southern Mississippi

Tim Kroft
Auburn University at Montgomery

Mary Rose Lamb
University of Puget Sound

Melanie Lee-Brown
Guilford College

Aime Levesque
University of Hartford

Joshua Loomis
Nova Southeastern University

Shawn Macauley
Muskegon Community College

William Mackay
Edinboro University of Pennsylvania

Cindy S. Malone
California State University at Northridge

Teresa McElhinny
Michigan State University

Karen McGinnis
Florida State University

Amy McMillan
SUNY Buffalo State

Steven Mezik
Herkimer County Community College

Brook Milligan
New Mexico State University

Marcie Moehnke
Baylor University

Charles Molnar
Camosun College

Jessica L. Moore
Western Carolina University

Sarah Mordan-McCombs
Franklin College of Indiana

Ashley Morris

Middle Tennessee State University

Cam Muir
University of Hawaii at Hilo

Karolina Mukhtar
University of Alabama at Birmingham

Elbert Myles
Tennessee State University

Todd Nickle
Mount Royal University

John Niedzwiecki
Belmont University

Selene Nikaido
University of Central Missouri

Margaret Olney
Saint Martin's University

Sally Pasion
San Francisco State University

Ann Paterson
Williams Baptist College

Helen Piontkivska
Kent State University

Uwe Pott
University of Wisconsin-Green Bay

Michael Lee Robinson
Miami University

Charles Sackerson
California State University

Stephanie Schroeder
Webster University

Rodney Scott
Wheaton College

Rebecca Seipelt-Thiemann
Middle Tennessee State University

Barkur S. Shastry
Oakland University

Mark Shotwell
Slippery Rock University

Wendy Shuttleworth
Lewis/Clark State College

Agnes Southgate
College of Charleston

Walter Sotero
University of Central Florida

Ron Strohmeyer
Northwest Nazarene University

David Thompson

Fresno Pacific University

Sandra Thompson-Jaeger
Christian Brothers University

Douglas Thrower
University of California at Santa Barbara

Kathleen Toedt
Housatonic Community College

Cynthia van Golen
Delaware State University

Nanette van Loon
Borough of Manhattan Community College

Sara Volk
Texas State University

Laurence von Kalm
University of Central Florida

Erik Vollbrecht
Iowa State University

Alan Waldman
University of South Carolina

Daniel Williams
Winston-Salem State University

Lise D. Wilson
Siena College

Kathleen Wood
University of Mary Hardin-Baylor

Lev Yampolsky
East Tennessee State University

Malcolm Zellars
Georgia State University

Ming Zheng
Gordon College

Sumário

Capítulo 1 Introdução à Genética

Albinismo nos hopis

1.1 A genética é importante para todos nós como indivíduos e sociedade, e para o estudo da Biologia

O papel da genética na Biologia

Diversidade genética e evolução

Divisões da genética

Organismos-modelo da genética

1.2 Os humanos praticam genética há milhares de anos

As primeiras técnicas genéticas e a compreensão da hereditariedade

O início da ciência genética

O futuro da genética

1.3 Alguns conceitos são fundamentais para o início de nossa jornada na Genética

Capítulo 2 Cromossomos e Reprodução Celular

O enigma dos cegos

2.1 As células procarióticas e eucarióticas apresentam diferenças em várias características genéticas

2.2 A reprodução celular exige a cópia do material genético, a separação das cópias e a divisão celular

Reprodução da célula procariótica

Reprodução da célula eucariótica

O ciclo celular e a mitose

Consequências genéticas do ciclo celular

2.3 A reprodução sexuada produz variação genética pela meiose

Meiose

Fontes de variação genética na meiose

Separação das cromátides-irmãs e dos cromossomos homólogos

Meiose nos ciclos de vida dos animais e das plantas

Capítulo 3 Princípios Básicos da Hereditariedade

A genética do cabelo ruivo

3.1 Gregor Mendel descobriu os princípios básicos da hereditariedade

O sucesso de Mendel

Terminologia genética

3.2 Os cruzamentos mono-híbridos revelam o princípio da segregação e o conceito de dominância

O que os cruzamentos mono-híbridos revelam

A natureza molecular dos alelos

Como prever os desfechos dos cruzamentos genéticos

O cruzamento-teste

Símbolos genéticos

3.3 Os cruzamentos di-híbridos revelam o princípio da segregação independente

Cruzamentos di-híbridos

Princípio da segregação independente

O princípio da segregação independente e a meiose

Como aplicar a probabilidade e o diagrama ramificado para os cruzamentos di-híbridos

O cruzamento-teste di-híbrido

3.4 As razões observadas da prole podem desviar-se das razões esperadas ao acaso

Teste qui-quadrado de adequação de ajuste

Capítulo 4 Sexo | Determinação e Características

Curiosidades sobre o sexo do ornitorrinco

4.1 O sexo é determinado por vários mecanismos diferentes

Sistemas cromossômicos para determinação do sexo

Determinação do sexo pelo gene

Determinação do sexo pelo ambiente

Determinação do sexo na *Drosophila melanogaster*

Determinação do sexo nos seres humanos

4.2 As características ligadas ao sexo são determinadas por genes nos cromossomos sexuais

Os olhos brancos ligados ao X na *Drosophila*

Não disjunção e teoria cromossômica da hereditariedade

Daltonismo ligado ao X nos seres humanos

Símbolos para os genes ligados ao X

Características ligadas ao Z

Características ligadas ao Y

4.3 A compensação da dosagem iguala a quantidade de proteína produzida por genes ligados ao X e genes autossômicos em alguns animais

Hipótese de Lyon

Mecanismo de inativação aleatória do X

Capítulo 5 Extensões e Modificações dos Princípios Básicos

A estranha genética do caracol canhoto

5.1 Fatores adicionais em um único *locus* podem afetar os resultados dos cruzamentos genéticos

Tipos de dominância

Penetrância e expressividade

Alelos letais

Alelos múltiplos

5.2 A interação genética ocorre quando os genes em múltiplos *loci* determinam um único fenótipo

A interação gênica e a produção de novos fenótipos

Interação gênica com epistasia

Complementação | Como determinar se as mutações estão no mesmo *locus* ou em *loci* diferentes

A complexa genética da cor do pelo nos cães

5.3 O sexo influencia a herança e a expressão dos genes de várias maneiras

Características influenciadas e características limitadas pelo sexo

Herança citoplasmática

Efeito genético materno

Imprinting genômico

5.4 A antecipação é a expressão mais forte ou precoce dos traços nas gerações seguintes

5.5 A expressão de um genótipo pode ser afetada por efeitos ambientais

Efeitos ambientais no fenótipo

A herança das características contínuas

Capítulo 6 Análise de Heredograma, Aplicações e Exame Genético

O mistério das impressões digitais desaparecidas

6.1 O estudo da genética nos seres humanos é limitado por características especiais da biologia e da cultura humanas

6.2 Os geneticistas usam frequentemente os heredogramas para estudar a hereditariedade das características nos seres humanos

Símbolos dos heredogramas

Análise dos heredogramas

Traços autossômicos recessivos

Traços autossômicos dominantes

Traços recessivos ligados ao X

Traços dominantes ligados ao X

Traços ligados ao Y

6.3 O estudo sobre gêmeos e adoções pode ajudar a avaliar a importância dos genes e do meio ambiente

Tipos de gêmeos

Concordância em gêmeos

Estudo da asma em gêmeos

Estudos sobre adoção

6.4 O aconselhamento e os exames genéticos fornecem informações às pessoas preocupadas com doenças e traços genéticos

Aconselhamento genético

Exame genético

Como interpretar os testes genéticos

Testes genéticos de coleta domiciliares

Discriminação genética e privacidade

Capítulo 7 Ligação Gênica, Recombinação e Mapeamento de Genes Eucarióticos

Os genes ligados e a calvície

7.1 Os genes ligados não se separam de forma independente

7.2 Os genes ligados se separam juntos, enquanto o *crossing over* produz recombinação entre eles

Notação dos cruzamentos com ligação

Ligação completa comparada com segregação independente

Crossing over com genes ligados

Como calcular a frequência de recombinação

Acoplamento e repulsão

Evidência para a base física da recombinação

Como prever os desfechos dos cruzamentos com genes ligados

Análise de segregação independente

Mapeamento de genes com frequências de recombinação

Como construir um mapa genético com o uso de cruzamentos-teste de dois pontos

7.3 Um cruzamento-teste de três pontos pode ser usado para mapeamento de três genes ligados

Como construir um mapa genético com o uso de cruzamento-teste de três pontos

Efeito dos *crossing overs* múltiplos

Como mapear os genes humanos

Mapeamento com marcadores moleculares

Genes podem ser localizados por estudos de associação genômica ampla

7.4 Os métodos de mapeamento físico são usados para determinar as posições físicas dos genes em cromossomos específicos

Hibridização de células somáticas

Mapeamento de deleção

Mapeamento cromossômico físico por meio de análise molecular

7.5 As taxas de recombinação exibem grande variação

Capítulo 8 Variação Cromossômica

Como criar bananas de melhor qualidade

8.1 Mutações cromossômicas incluem rearranjos aneuploides e poliploides

Morfologia do cromossomo

Tipos de mutações cromossômicas

8.2 Rearranjos cromossômicos modificam a estrutura dos cromossomos

Duplicações

Deleções

Inversões

Translocações

Sítios frágeis

Variações no número de cópias

8.3 Aneuploidia é o aumento ou a redução no número de cromossomos individuais

Tipos de aneuploidia

Efeitos da aneuploidia

Aneuploidia nos seres humanos

Dissomia uniparental

Mosaicismo

8.4 Poliploidia é a ocorrência de mais de dois conjuntos de cromossomos

Autopoliploidia

Alopoliploidia

Importância da poliploidia

Capítulo 9 Sistemas Genéticos das Bactérias e dos Vírus

A vida no mundo das bactérias

9.1 A análise genética das bactérias exige métodos especiais

Diversidade bacteriana

Técnicas para o estudo das bactérias

O genoma bacteriano

Plasmídios

9.2 As bactérias trocam os genes por meio de conjugação, transformação e transdução

Conjugação

Transferência natural dos genes e resistência a antibióticos

Transformação em bactérias

Sequências de genoma bacteriano

Transferência horizontal de genes

9.3 Vírus são sistemas simples, replicantes, receptivos à análise genética

Técnicas para o estudo dos bacteriófagos

Transdução | Como usar os fagos para mapear os genes bacterianos

Mapeamento dos genes nos fagos

Análise intraestrutural dos genes dos bacteriófagos

Vírus com RNA

Vírus da imunodeficiência humana e AIDS

Influenza | Gripe

Capítulo 10 DNA | A Natureza Química do Gene

Viagens pelo Ártico e DNA antigo

10.1 O material genético tem várias características importantes

10.2 Todas as informações genéticas são codificadas na estrutura do DNA ou RNA

Estudos iniciais sobre o DNA

DNA como fonte de informação genética

Descoberta da estrutura tridimensional do DNA por Watson e Crick

RNA como material genético

10.3 O DNA consiste em duas cadeias de nucleotídios complementares e antiparalelas que formam uma dupla-hélice

Estrutura primária do DNA

Estruturas secundárias do DNA

10.4 Estruturas especiais podem se formar no DNA e no RNA

Capítulo 11 Estrutura do Cromossomo e DNA das Organelas

Telômeros e adversidade na infância

11.1 Muito DNA é condensado em uma célula

- Superenrolamento
- Cromossomo bacteriano
- Cromossomos eucarióticos
- Mudanças na estrutura da cromatina

11.2 Os cromossomos eucarióticos têm centrômeros e telômeros

- Estrutura do centrômero
- Estrutura do telômero

11.3 O DNA eucariótico contém as várias classes de variação de sequências

- Desnaturação e renaturação do DNA
- Tipos de sequências de DNA nos eucariotos

11.4 O DNA de organelas tem características únicas

- Estrutura da mitocôndria e do cloroplasto
- Teoria endossimbiótica
- Herança uniparental de traços codificados pelas organelas
- Genoma mitocondrial
- Evolução do DNA mitocondrial
- Dano ao DNA mitocondrial e sua associação ao envelhecimento
- Genoma de cloroplasto
- Transmissão da informação genética entre os genomas nucleares, mitocondriais e cloroplásticos ao longo da evolução

Capítulo 12 Replicação e Recombinação de DNA

Topoisomerase, replicação e câncer

12.1 As informações genéticas precisam ser copiadas com exatidão a cada divisão celular

12.2 Toda replicação de DNA ocorre de modo semiconservativo

- Experimento de Meselson e Stahl
- Modos de replicação
- Exigências da replicação
- Sentido de replicação

12.3 A replicação bacteriana exige numerosas enzimas e proteínas

- Iniciação
- Desenrolamento
- Alongamento
- Término
- Fidelidade da replicação do DNA

12.4 A replicação do DNA eucariótico é semelhante à replicação bacteriana, mas existem algumas diferenças

- Origens eucarióticas
- Licenciamento da replicação do DNA
- Desenrolamento
- DNA polimerases eucarióticas

Montagem do nucleossomo
Localização da replicação no núcleo
Síntese de DNA e ciclo celular
Replicação nas extremidades dos cromossomos
Replicação em archaea

12.5 Recombinação consiste em ruptura, alinhamento e reparo das fitas de DNA
Modelos de recombinação
Enzimas necessárias à recombinação
Conversão gênica

Capítulo 13 Transcrição

Envenenamento por cicuta-verde

13.1 O RNA, composto por uma fita única de ribonucleotídios, participa em várias funções celulares
Um mundo primitivo de RNA
Estrutura do RNA
Classes de RNA

13.2 Transcrição é a síntese de uma molécula de RNA a partir de um molde de DNA
Molde
Substrato para transcrição
Aparato de transcrição

13.3 A transcrição bacteriana consiste em iniciação, alongamento e término
Iniciação
Alongamento
Término

13.4 A transcrição eucariótica é semelhante à bacteriana, mas tem diferenças importantes
Transcrição e estrutura do nucleossomo
Promotores
Iniciação
Alongamento
Término

13.5 A transcrição em Archaea é mais semelhante à transcrição nos eucariotos do que à transcrição nas eubactérias

Capítulo 14 Moléculas de RNA e Processamento do RNA

Uma doença na família real

14.1 Muitos genes têm estruturas complexas
Organização dos genes
Íntrons
Conceito de gene revisitado

14.2 Os RNAs mensageiros que codificam as sequências de aminoácidos das proteínas são modificados após a transcrição nos eucariotos
Estrutura do RNA mensageiro

Processamento do pré-mRNA
Adição do cap 5′
Adição da cauda poli(A)
Splicing ou recomposição do RNA
Vias de processamento alternativas
Edição do RNA

14.3 Os RNAs transportadores, ligados aos aminoácidos, são modificados após a transcrição nas células bacterianas e eucarióticas
Estrutura do RNA transportador
Estrutura e processamento do gene do RNA transportador

14.4 O RNA ribossômico, um componente do ribossomo, também é processado após a transcrição
A estrutura do ribossomo
Estrutura e processamento do gene do RNA ribossômico

14.5 Pequenas moléculas de RNA participam em várias funções
Interferência por RNA
Pequenos RNAs de interferência e microRNAs
RNAs que interagem com Piwi
CRISPR RNA

14.6 Os RNAs longos não codificadores regulam a expressão gênica

Capítulo 15 Código Genético e Tradução

Hutterites, ribossomos e a síndrome de Bowen-Conradi

15.1 Muitos genes codificam proteínas
Hipótese de um gene, uma enzima
Estrutura e função das proteínas

15.2 O código genético determina como a sequência de nucleotídios especifica a sequência de aminoácidos de uma proteína
Como decifrar o código genético
Degeneração do código
Quadro de leitura e códons de iniciação
Códons de terminação
Universalidade do código

15.3 Aminoácidos são montados em uma proteína por meio da tradução
Ligação dos aminoácidos aos RNAs transportadores
Iniciação da tradução
Alongamento
Terminação

15.4 Propriedades adicionais do RNA e dos ribossomos influenciam a síntese de proteínas
Estrutura tridimensional do ribossomo
Polirribossomos
Vigilância do RNA mensageiro
Dobradura e modificações pós-tradução das proteínas

Tradução e antibióticos

Capítulo 16 Controle da Expressão Gênica nas Bactérias

Óperons e a célula de ruído

16.1 A regulação da expressão gênica é crítica para todos os organismos

Genes e elementos regulatórios

Níveis de regulação gênica

Proteínas ligadoras do DNA

16.2 Os óperons controlam a transcrição nas células bacterianas

Estrutura do óperon

Controles negativo e positivo | Óperons induzíveis e reprimíveis

Óperon *lac* de *E. coli*

Mutações *lac*

Controle positivo e repressão catabólica

Óperon *trp* de *E. coli*

Acentuadores bacterianos

16.3 Alguns óperons regulam a transcrição por meio da atenuação, o término prematuro da transcrição

Atenuação no óperon *trp* da *E. coli*

Por que ocorre atenuação no óperon *trp*?

16.4 As moléculas de RNA controlam a expressão em alguns genes bacterianos

RNA antissenso

Riboswitches

Repressão mediada do RNA por meio de ribozimas

Capítulo 17 Controle da Expressão Gênica nos Eucariotos

As diferenças genéticas que nos tornam humanos

17.1 Células eucarióticas e bactérias têm muitas características de regulação gênica em comum, mas existem diferenças importantes

17.2 Alterações na estrutura da cromatina influenciam a expressão dos genes

Hipersensibilidade da DNase I

Remodelagem da cromatina

Modificação das histonas

Metilação do DNA

17.3 O início da transcrição é regulado por fatores de transcrição e proteínas reguladoras

Ativadores e coativadores da transcrição

Repressores da transcrição

Acentuadores e insuladores

Regulação da parada e alongamento transcricional

Regulação gênica coordenada

17.4 Alguns genes são regulados por processamento e degradação do RNA

Regulação gênica por meio da recomposição do RNA

Degradação do RNA

17.5 A interferência por RNA é um mecanismo importante de regulação gênica
RNAs de interferência pequenos e microRNAs
Mecanismos de regulação gênica pela interferência por RNA
Controle do desenvolvimento por interferência por RNA

17.6 Alguns genes são regulados por processos que afetam a tradução ou por modificações das proteínas

Capítulo 18 Mutações do Gene e Reparo do DNA

Uma mosca sem coração

18.1 As mutações são alterações herdadas na sequência do DNA
Importância das mutações
Categorias de mutações
Tipos de mutações gênicas
Efeitos fenotípicos das mutações
Mutações supressoras
Taxas de mutação

18.2 As mutações são potencialmente causadas por vários fatores diferentes
Erros espontâneos de replicação
Mudanças químicas espontâneas
Mutações quimicamente induzidas
Radiação

18.3 As mutações são objeto de intenso estudo pelos geneticistas
Como detectar as mutações com o teste de Ames
Exposição de seres humanos à radiação

18.4 Os elementos t de transposição provocam mutações
Características gerais dos elementos de transposição
Transposição
Efeitos mutagênicos da transposição
Elementos de transposição nas bactérias
Elementos de transposição nos eucariotos
Elementos de transposição são importantes para a evolução humana

18.5 Várias vias reparam as mudanças no DNA
Reparo de pareamento errado
Reparo direto
Reparo por excisão de bases
Reparo por excisão de nucleotídio
Reparo de quebras de fita dupla
DNA polimerases translesão
Doenças genéticas e reparo do DNA defeituoso

Capítulo 19 Análise da Genética Molecular e Biotecnologia

Como ajudar o cego a enxergar

19.1 As técnicas de genética molecular revolucionaram a biologia

Revolução da genética molecular
Como trabalhar no nível molecular

19.2 Técnicas moleculares são usadas para isolar, recombinar e ampliar os genes
Corte e união dos fragmentos de DNA
Visualização dos fragmentos de DNA
Localização dos fragmentos de DNA com *Southern blotting* e sondas
Clonagem de genes
Aplicação | Engenharia genética das plantas com pesticidas
Amplificação dos fragmentos de DNA com reação em cadeia da polimerase

19.3 As técnicas moleculares podem ser usadas para encontrar genes de interesse
Bibliotecas gênicas
Hibridização *in situ*
Clonagem posicional
Aplicação | Isolamento do gene para a fibrose cística

19.4 As sequências podem ser determinadas e analisadas
Polimorfismos de comprimento do fragmento de restrição
Sequenciamento de DNA
Tecnologias de sequenciamento de próxima geração
Fingerprinting do DNA
Aplicação | Identificação dos corpos no colapso das torres do World Trade Center

19.5 Técnicas moleculares são usadas para analisar a função dos genes
Genética direta e reversa
Criação de mutações aleatórias
Mutagênese direcionada pelo sítio
Animais transgênicos
Camundongos knockout
Como silenciar os genes com RNAi
Aplicação | Uso do RNAi para o tratamento de doenças humanas

19.6 A biotecnologia utiliza o poder da genética molecular
Produtos farmacêuticos
Bactérias especializadas
Produtos agrícolas
Exame genético
Terapia gênica

Capítulo 20 Genômica e Proteômica

Como decifrar a dança das abelhas | O genoma da abelha

20.1 A genômica estrutural determina as sequências de DNA de genomas inteiros
Mapas genéticos
Mapas físicos
Como sequenciar um genoma inteiro
Projeto Genoma Humano
Polimorfismos de único nucleotídio

Variações no número de cópias
Sítios de sequências marcadas e etiquetas de sequência expressa
Bioinformática
Metagenômica
Biologia sintética

20.2 A genômica funcional determina a função dos genes ao usar abordagens com base na genômica
Como prever a função a partir da sequência
Expressão gênica e microarranjos
Expressão gênica e sequências repórter
Mutagênese genômica global

20.3 A genômica comparativa estuda como os genomas evoluem
Genomas procarióticos
Genomas eucarióticos
Genômica comparativa da *Drosophila*
Genoma humano

20.4 A proteômica analisa o conjunto complexo de proteínas encontrado em uma célula
Determinação das proteínas celulares
Captura por afinidade
Microarranjo de proteínas
Proteômica estrutural

Capítulo 21 Epigenética

Como a dieta do seu avô pode afetar a sua saúde

21.1 O que é epigenética?

21.2 Vários processos moleculares levam a mudanças epigenéticas
Metilação do DNA
Modificações da histona
Efeitos epigenéticos produzidos pelas moléculas de RNA

21.3 Os processos epigenéticos produzem um conjunto variado de efeitos
Paramutação
Epigenética comportamental
Efeitos epigenéticos das substâncias químicas do ambiente
Efeitos epigenéticos transgeração sobre o metabolismo
Efeitos epigenéticos em gêmeos monozigóticos
Inativação do X
Alterações epigenéticas associadas à diferenciação celular
Imprinting genômico

21.4 Epigenoma

Capítulo 22 Genética Desenvolvimental e Imunogenética

Origem dos esgana-gatas sem espinhas dorsais e pélvicas

22.1 O desenvolvimento se dá por diferenciação celular

Experimentos de clonagem em plantas
Experimentos de clonagem em animais

22.2 A formação do padrão em *Drosophila* serve como modelo para o controle genético do desenvolvimento
Desenvolvimento da mosca-da-fruta
Genes da polaridade do ovo
Genes de segmentação
Genes homeóticos em *Drosophila*
Genes *homeobox* em outros organismos
Mudanças epigenéticas no desenvolvimento

22.3 Os genes controlam o desenvolvimento das flores nas plantas
Anatomia da flor
Controle genético de desenvolvimento da flor

22.4 A morte celular programada é uma parte integral do desenvolvimento

22.5 O estudo do desenvolvimento revela padrões e processos de evolução

22.6 O desenvolvimento da imunidade pelo rearranjo genético
Organização do sistema imunológico
Estrutura da imunoglobulina
Geração da diversidade dos anticorpos
Diversidade do receptor de linfócito T
Genes do complexo principal de histocompatibilidade
Genes e transplantes de órgãos

Capítulo 23 A Genética do Câncer

A paladina e a disseminação do câncer

23.1 O câncer é um grupo de doenças caracterizado pela proliferação celular
Formação do tumor
Câncer como uma doença genética
Papel dos fatores ambientais no câncer

23.2 Mutações em vários tipos diferentes de genes contribuem para o câncer
Oncogenes e genes supressores de tumor
Mutações nos genes que controlam o ciclo de divisão celular
Genes de reparo do DNA
Genes que regulam a telomerase
Genes que promovem a vascularização e a disseminação dos tumores
MicroRNAs e câncer
Projetos de genoma do câncer

23.3 As mudanças epigenéticas estão frequentemente associadas com o câncer

23.4 O câncer colorretal surge em decorrência de mutação sequencial de vários genes

23.5 Mudanças no número e na estrutura do cromossomo estão frequentemente associadas com câncer

23.6 Os vírus estão associados com alguns cânceres

Capítulo 24 Genética Quantitativa

Óleo de milho e genética quantitativa

24.1 As características quantitativas variam continuamente e muitas são influenciadas por alelos em múltiplos *loci*

Relação entre genótipo e fenótipo

Tipos de características quantitativas

Herança poligênica

Cor do grão do trigo

Como determinar o número de genes para uma característica poligênica

24.2 Métodos estatísticos são cruciais para analisar características quantitativas

Distribuições

Amostras e populações

Média

Variância e desvio padrão

Correlação

Regressão

Como aplicar a estatística no estudo de uma característica poligênica

24.3 A herdabilidade é usada para estimar a proporção de variação em um traço genético

Variância fenotípica

Tipos de herdabilidade

Como calcular a herdabilidade

Limitações da herdabilidade

Como localizar os genes que afetam as características quantitativas

24.4 Os traços geneticamente variáveis mudam em resposta à seleção

Como prever a resposta à seleção

Limites para a resposta à seleção

Respostas correlacionadas

Capítulo 25 Genética de Populações

Resgate genético do carneiro-das-montanhas

25.1 As frequências genotípicas e alélicas são usadas para descrever o *pool* de genes de uma população

Como calcular as frequências genotípicas

Como calcular as frequências alélicas

25.2 A lei de Hardy-Weinberg descreve o efeito da reprodução nas frequências genotípicas e alélicas

Frequências genotípicas no equilíbrio de Hardy-Weinberg

Exame mais detalhado da lei de Hardy-Weinberg

Implicações da lei de Hardy-Weinberg

Extensões da lei de Hardy-Weinberg

Teste para proporções de Hardy-Weinberg

Como estimar as frequências alélicas com a lei de Hardy-Weinberg

25.3 O acasalamento não aleatório afeta as frequências genotípicas de uma população

25.4 Várias forças evolutivas causam mudanças nas frequências alélicas

Mutação
Migração
Deriva genética
Seleção natural

Capítulo 26 Genética Evolutiva

Os genes do paladar nos macacos que cospem

26.1 A evolução ocorre por mudança genética nas populações

26.2 Muitas populações naturais apresentam níveis elevados de variação genética
Variação molecular
Variação de proteína
Variação na sequência de DNA

26.3 Novas espécies surgem pela evolução do isolamento reprodutivo
Conceito biológico de espécie
Mecanismos de isolamento reprodutivo
Modos de especiação
Diferenciação genética associada com a especiação

26.4 A história evolutiva de um grupo de organismos pode ser reconstruída pelo estudo das mudanças nas características homólogas
Alinhamento de sequências homólogas
Construção das árvores filogenéticas

26.5 Os padrões de evolução são revelados por mudanças moleculares
Taxas de evolução molecular
Relógio molecular
Evolução por meio de mudanças na regulação gênica
Evolução do genoma

Guia de Referência para Organismos-Modelo da Genética
Mosca-da-fruta *Drosophila melanogaster*
Bactéria E*scherichia coli*
Nematódeo *Caenorhabditis elegans*
Planta *Arabidopsis thaliana*
Camundongo *Mus musculus*
Levedura *Saccharomyces cerevisiae*

Glossário

Respostas a Questões e Problemas Selecionados

1 Introdução à Genética

Albinismo nos hopis

A mil pés acima do deserto, Black Mesa domina o horizonte do Enchanted Desert e fornece um ponto de referência familiar para os viajantes que passam pela região nordeste do Arizona. Black Mesa não é somente uma característica geológica evidente, mas também – e mais importante – é o lar ancestral dos nativos americanos hopis. As extremidades digitiformes de Mesa avançam para o deserto, e, ao longo ou no topo de cada uma delas, existe uma vila hopi. A maior parte das vilas é bem pequena, com apenas algumas dezenas habitantes, mas eles são muito idosos. A vila Oraibi existe em Black Mesa desde 1150 d.C., sendo o mais antigo assentamento continuamente ocupado da América do Norte.

Em 1900, Alêes Hrdliêeka, um antropólogo e médico que trabalhava para o American Museum of Natural History, visitou as vilas hopis de Black Mesa e relatou uma descoberta impressionante. Ele encontrou 11 pessoas de pele branca, não caucasianos, entre os hopis, que de fato eram nativos americanos hopi de pele branca. Eles eram portadores de uma condição genética conhecida como albinismo.

O albinismo é provocado por um defeito em uma das enzimas necessárias para produzir melanina, o pigmento que escurece nossa pele, pelos e olhos. Pessoas com albinismo não produzem melanina ou produzem pequenas quantidades e, consequentemente, têm cabelo branco, pele clara e nenhum pigmento na íris dos olhos. A melanina normalmente protege o DNA das células da pele dos efeitos nocivos da radiação ultravioleta da luz do sol, e a existência desse composto no olho em desenvolvimento é essencial para a visão adequada.

A base genética do albinismo foi descoberta pelo médico inglês Archibald Garrod, que identificou, em 1908, que a condição era hereditária, como uma característica autossômica recessiva, o que significa que uma pessoa tem de receber duas cópias de uma mutação albina – uma do pai e uma da mãe – para apresentar o albinismo. A natureza molecular das mutações que levam ao albinismo foi elucidada recentemente. O albinismo nos seres humanos é causado por defeitos em um dos vários genes diferentes que controlam a síntese e o armazenamento da melanina; podem ocorrer muitos tipos diferentes de mutações em cada gene, cada qual podendo levar ao albinismo. É mais provável que a forma de albinismo encontrada nos hopis seja o albinismo oculocutâneo (que afeta os olhos e a pele), tipo II, devido ao defeito no gene *OCA2* no cromossomo 15.

Os hopis não são os únicos a terem albinos entre os membros de sua tribo. O albinismo é encontrado em quase todos os grupos étnicos humanos e está registrado em escritos antigos, sendo provável que exista desde o primórdio da humanidade. O que torna os hopis únicos é a alta frequência de albinismo na sua população. Ele é raro na maior parte dos grupos humanos, ocorre em cerca de 1 em 20.000 pessoas, mas alcança uma frequência de 1 em 200 nos vilarejos da Black Mesa, 100 vezes mais frequente que nas outras populações.

Por que o albinismo é tão frequente entre os hopis? A resposta a essa pergunta não é completamente conhecida, mas os geneticistas que estudam albinismo nos hopis especulam que a alta frequência do gene albino esteja relacionada com o local especial que o albinismo ocupou na cultura hopi. Por grande parte da sua história, os hopis consideram os membros da tribo com albinismo como sendo importantes e especiais. As pessoas com albinismo eram consideradas bonitas, limpas e inteligentes. Ter várias pessoas com albinismo em uma vila era considerado um bom sinal, um símbolo de que as pessoas da vila tinham o puro sangue hopi. Os albinos realizavam cerimônias hopis e detinham posições de liderança dentro da tribo, como chefes, curandeiros e líderes religiosos.

Os hopis albinos também recebem tratamento especial nas atividades diárias. Os hopis cultivam pequenas hortas no sopé da Black Mesa há séculos. Diariamente, na estação de plantação, os homens da tribo viajam para a base da Black Mesa e gastam boa parte do dia sob a luz do sol do sudoeste cuidando do milho e vegetais. Com pouca ou nenhuma melanina na sua pele, as pessoas com albinismo são muito suscetíveis à queimadura pelo sol e apresentam maior incidência de câncer de pele quando expostos ao sol. Além disso, eles não enxergam bem na luz do dia. Portanto, os homens hopis com albinismo eram liberados dessa atividade masculina e podiam permanecer na vila com as mulheres da tribo, realizando outras atividades.

Durante a estação de plantação, os homens albinos eram os únicos homens da tribo na vila com as mulheres durante o dia e, portanto, eles desfrutavam de uma vantagem na reprodução, o que ajudou a espalhar seus genes do albinismo. Além disso, as vantagens especiais dadas aos albinos permitiram que eles evitassem os efeitos prejudiciais do albinismo – aumento dos casos de câncer de pele e visão ruim. O tamanho pequeno da tribo provavelmente também teve seu papel ao possibilitar o aumento da frequência do gene albino. Independentemente dos fatores que levaram à alta frequência de albinismo, está claro que os hopis respeitavam e valorizavam os membros da sua tribo com essa característica específica. Infelizmente, as pessoas com condições genéticas, em muitas sociedades, estão frequentemente sujeitas a discriminação e preconceito. **Resolva os Problemas 1 e 25**

A genética é um dos campos da ciência que avança mais rapidamente, com novas descobertas importantes descritas todos os meses. Ao ler os principais jornais e revistas, você verá artigos relacionados com a genética: a finalização de outro genoma, como o da borboleta Monarca; a descoberta de genes que afetam doenças importantes, incluindo esclerose múltipla, depressão e câncer; um relato de DNA analisado a partir de animais extintos como o mamute-lanoso; e a identificação de genes que afetam a pigmentação da pele, a altura e a capacidade de aprendizado nos humanos. Até entre anúncios, é provável que você encontre testes genéticos: para determinar a ascendência, paternidade e suscetibilidade das pessoas para doenças e distúrbios. Estas novas descobertas e usos da genética têm implicações econômicas e éticas importantes, tornando o estudo da genética relevante, oportuno e interessante.

Este capítulo apresenta a genética e revisa alguns conceitos que você viu rapidamente em um curso de Biologia. Começamos por considerar a importância da genética para cada um de nós, a sociedade como um todo e os estudantes de Biologia. Então nos voltamos para a história da genética e como este campo se desenvolveu. O final do capítulo apresenta alguns termos fundamentais e os princípios da genética que são usados neste livro.

1.1 A genética é importante para todos nós como indivíduos e sociedade, e para o estudo da Biologia

O albinismo entre os hopis ilustra o papel importante que os genes desempenham em nossas vidas. Este defeito genético, entre os 20.000 genes humanos, altera completamente a vida de um hopi. Ele altera sua função, papel na sociedade hopi e as relações com os outros membros da tribo. Todos temos genes que influenciam nossas vidas de forma significativa. Os genes alteram nossa altura, peso, cor do cabelo e pigmentação da pele. Eles afetam nossa suscetibilidade a muitas doenças e distúrbios (**Figura 1.1**) e até contribuem para nossa inteligência e personalidade. Os genes são fundamentais para quem e o que somos.

Embora a ciência da genética seja relativamente nova, se comparada com outras, como Astronomia e Química, as pessoas compreendem a natureza hereditária das características e exercem a genética há milhares de anos. O surgimento da agricultura começou quando as pessoas começaram a aplicar os princípios da genética à domesticação de plantas e animais. Hoje, as principais culturas e animais usados na agricultura são muito diferentes dos seus progenitores selvagens, tendo sofrido grandes alterações genéticas que aumentaram seu rendimento e forneceram várias características desejáveis, como resistência a doenças e pragas, qualidades nutricionais especiais e características que facilitam a coleta. A Revolução Verde, que expandiu a produção de alimentos pelo mundo entre as décadas de 1950 e 1960, apoiou-se muito na aplicação da genética. Hoje, o milho, a soja e outras culturas geneticamente modificadas constituem uma proporção significativa de todos os alimentos produzidos no mundo.

A indústria farmacêutica é outra área na qual a genética tem um papel importante. Vários fármacos e aditivos alimentares são sintetizados a partir de fungos e bactérias que foram geneticamente manipulados para torná-los produtores eficientes destas substâncias. A indústria da biotecnologia usa as técnicas de genética molecular para desenvolver e produzir em grande quantidade substâncias de importância comercial. Os hormônios de crescimento, a insulina, os fatores de coagulação, as enzimas, os antibióticos, as vacinas e vários fármacos são comercialmente produzidos a partir de bactérias e outras células geneticamente modificadas. Ela também é usada para produzir bactérias que removem minerais do minério bruto, degradam substâncias químicas tóxicas e inibem a formação da geada prejudicial nas plantas da colheita.

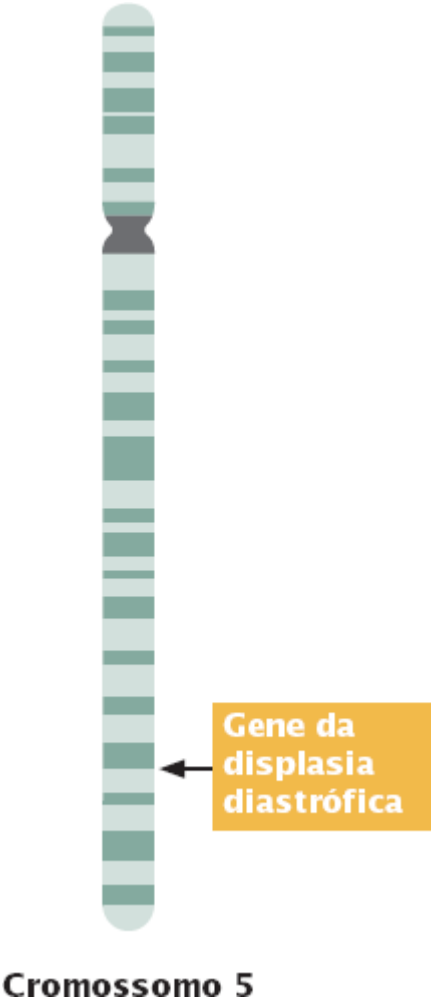


Figura 1.1 Os genes influenciam a suscetibilidade a muitas doenças e distúrbios. A displasia diastrófica é um distúrbio hereditário do crescimento que resulta em ossos e membros curtos e deformidades nas mãos. Esse distúrbio é causado por um defeito no gene *SLC26A2* no cromossomo 5.

A genética também é importante na Medicina. Os médicos reconhecem que muitas doenças e distúrbios têm um componente hereditário, incluindo distúrbios genéticos raros como a anemia falciforme e a doença de Huntington, assim como muitas doenças comuns como asma, diabetes e hipertensão. Os avanços na genética resultaram em revelações importantes sobre a natureza de doenças como câncer e no desenvolvimento de exames complementares, incluindo os que identificam patógenos e genes defeituosos. A terapia gênica – a alteração direta dos genes para tratar doenças humanas – é atualmente administrada a milhares de pacientes, embora seu uso ainda seja experimental e limitado a tratar alguns distúrbios.

O papel da genética na Biologia

Embora seja importante que todos entendam a genética, ela é crítica para o estudante de Biologia. A genética apresenta um dos princípios unificadores da Biologia: todos os organismos usam sistemas genéticos que têm várias características em comum. A genética também embasa o estudo de muitas outras disciplinas na Biologia. A evolução, por exemplo, é a mudança genética que ocorre com o passar do tempo; desse modo, o estudo da evolução requer uma compreensão da genética.

A Biologia desenvolvimental depende muito da genética: os tecidos e órgãos se desenvolvem pela expressão regulada dos genes (**Figura 1.2**). Até campos como taxonomia, ecologia e comportamento animal estão usando cada vez mais os métodos genéticos. O estudo de qualquer campo da Biologia ou Medicina não é completo sem uma compreensão perfeita dos genes e métodos genéticos.

Diversidade genética e evolução

A vida na Terra existe em um arranjo incrível de formas e características em quase todos os ambientes concebíveis. A vida também se caracteriza pela adaptação: muitos organismos estão primorosamente adaptados ao ambiente onde se encontram. A história da vida é uma crônica de novas formas de vida emergindo, formas antigas desaparecendo e formas existentes mudando.

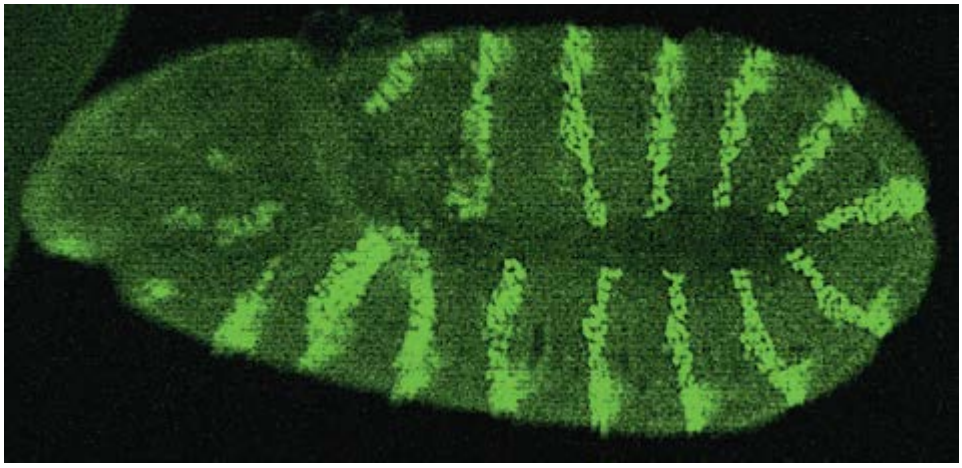


Figura 1.2 A chave para o desenvolvimento reside na regulação da expressão do gene. Este embrião precoce de mosca-da-fruta ilustra a expressão localizada do *gene engrailed*, que ajuda a determinar o desenvolvimento de segmentos do corpo na mosca adulta. (Stephen Paddock.)

Apesar da incrível diversidade, os organismos vivos têm uma característica importante em comum: todos usam sistemas genéticos semelhantes. Um conjunto completo de instruções genéticas para qualquer organismo é seu **genoma** e todos os genomas estão codificados nos ácidos nucleicos – DNA ou RNA. O sistema de codificação para as informações genômicas também é comum a todas as vidas: as instruções genéticas estão no mesmo formato, e, com raras exceções, as palavras do código são idênticas. Da mesma forma, o processo pelo qual as informações genéticas são copiadas e decodificadas é muito semelhante para todas as formas de vida. Essas características comuns da hereditariedade sugerem que toda a vida na Terra evoluiu do mesmo ancestral primordial que surgiu entre 3,5 bilhões e 4 bilhões de anos atrás. O biólogo Richard Dawkins descreve a vida como um rio de DNA que corre pelo tempo, conectando todos os organismos no passado e no presente.

Todos os organismos têm sistemas genéticos semelhantes, o que significa que o estudo dos genes de um organismo revela os princípios que se aplicam aos outros organismos. As investigações de como o DNA bacteriano é copiado (replicado), por exemplo, fornecem informações que se aplicam à replicação do DNA humano. Isto também significa que os genes funcionarão em células estranhas, o que torna a engenharia genética possível. Infelizmente, estes sistemas genéticos semelhantes também são a base para doenças como a AIDS (síndrome da imunodeficiência humana adquirida), na qual os genes virais são capazes de funcionar – às vezes com eficiência espantosa – nas células humanas.

A diversidade e a adaptação da vida são produto da evolução, que é simplesmente a mudança genética com o tempo. Ela é um processo de duas etapas: primeiro, as diferenças herdadas surgem de forma aleatória e, então, a proporção de indivíduos com diferenças específicas aumenta ou diminui. A variação genética é, portanto, a fundação de todas as alterações evolutivas e a base de todas as formas de vida que conhecemos. Além disso, técnicas de genética molecular são usadas na rotina para decifrar as relações evolutivas entre os organismos. Por exemplo, a análise recente do DNA isolado de fósseis de Neandertal forneceu novas informações sobre a relação entre eles e os humanos modernos, demonstrando que é provável que os Neandertais e os ancestrais dos humanos modernos tenham cruzado há cerca de 30.000 a 40.000 anos. A genética, o estudo da variação genética, é crítica para compreender o passado, o presente e o futuro da vida. **Resolva o Problema 17**

Conceitos

A hereditariedade afeta muitas das nossas características físicas, assim como nossa suscetibilidade para várias doenças e distúrbios. A genética contribui para os avanços na agricultura, indústria farmacêutica e Medicina, e é fundamental para a Biologia moderna. Todos os organismos usam sistemas genéticos semelhantes e a variação genética é o alicerce da diversidade de todas as formas de vida.

Checagem dos conceitos 1

Quais são algumas das implicações de todos os organismos terem sistemas genéticos semelhantes?

a. Que todas as formas de vida são geneticamente relacionadas.
b. Que as descobertas de pesquisa na função do gene de um organismo podem ser aplicadas a outros organismos.
c. Que os genes de um organismo podem, com frequência, existir e prosperar em outro organismo.
d. Todas as opções anteriores.

Divisões da genética

O estudo da genética consiste em três subdisciplinas principais: genética da transmissão, genética molecular e genética populacional (**Figura 1.3**). Também conhecida como genética clássica, a **genética da transmissão** inclui os princípios básicos da hereditariedade e como as características são transmitidas de uma geração para outra. Esta área se volta para a relação entre os cromossomos e a hereditariedade, a organização dos genes nos cromossomos e o mapeamento dos genes. Aqui, o foco é a organização individual – como um organismo herda sua composição genética e transmite seus genes para a próxima geração.

A **genética molecular** estuda a natureza química do próprio gene: como as informações genéticas são codificadas, replicadas e expressas. Ela inclui os processos celulares de replicação, transcrição e tradução (por meio da qual as informações genéticas são transferidas de uma molécula para outra) e a regulação dos genes (processo que controla a expressão das informações genéticas). O foco na genética molecular é o gene, sua estrutura, organização e função.

A **genética populacional** explora a composição genética dos grupos de membros da mesma espécie (populações) e como essa composição muda com a geografia e o passar do tempo. Como a evolução é uma mudança genética, a genética populacional é fundamentalmente o estudo da evolução. O foco da genética populacional é o grupo de genes encontrado em uma população.

A divisão do estudo da genética nesses três grupos é conveniente e tradicional, mas devemos admitir que os campos se sobrepõem e que cada subdivisão principal pode ser dividida em vários mais especializados, como genética dos cromossomos, genética bioquímica, genética quantitativa e assim por diante. Por outro lado, a genética pode ser subdividida por organismo (genética da mosca-da-fruta, milho ou bactérias) e cada um desses organismos pode ser estudado no nível de genética da transmissão, molecular e populacional. A genética moderna é um campo muito extenso, incluindo muitas subdisciplinas e especializações inter-relacionadas. Resolva o Problema 18

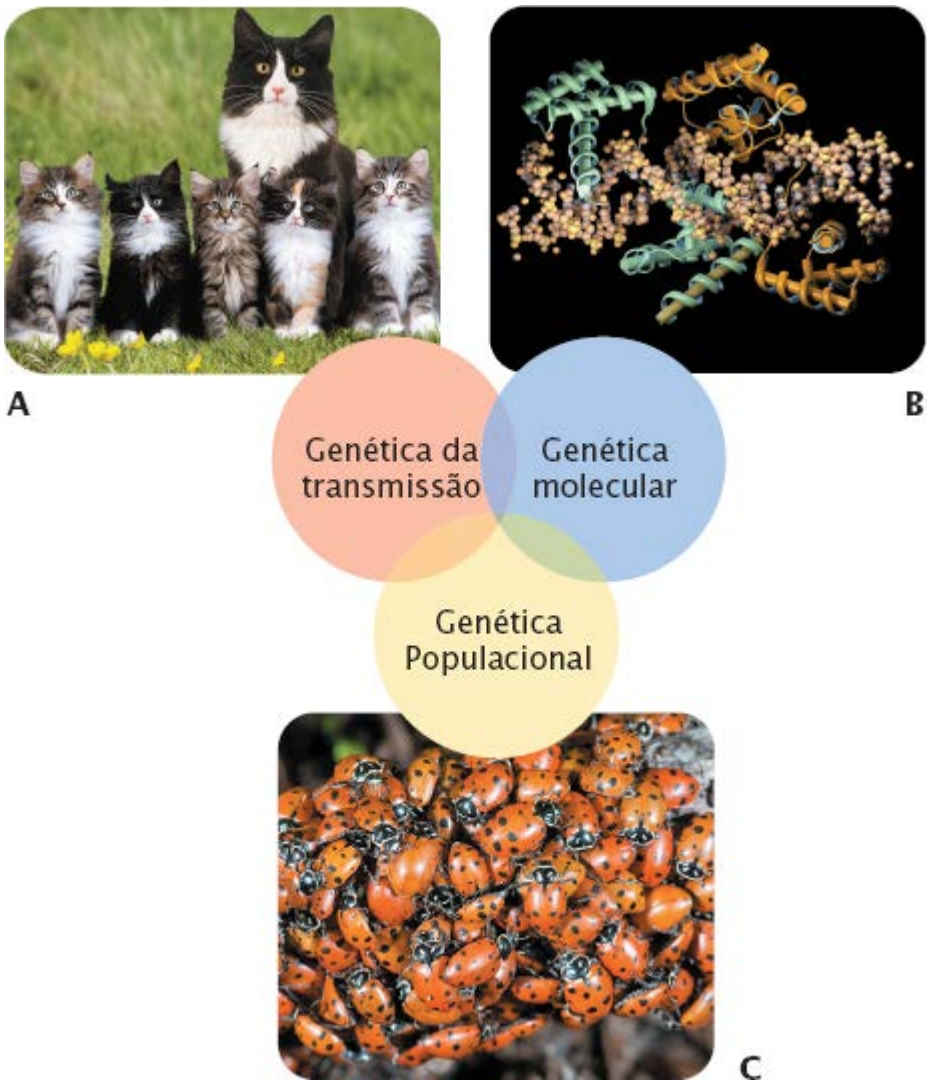


Figura 1.3 A genética pode ser subdividida em três campos inter-relacionados. (*Topo à esquerda*: Juniors Bildarchive/Alamy. *Topo à direita*: Martin McCarthy/Getty Images. *Embaixo*: Stuart Wilson/Science Source.)

Organismos-modelo da genética

Durante anos, os estudos de genética foram conduzidos em milhares de espécies diferentes, incluindo quase todos os principais grupos de bactérias, fungos, protistas, plantas e animais. Contudo, algumas espécies surgiram como **organismos-modelo da genética** – organismos com características que os tornam particularmente úteis para análise genética e sobre os quais existe uma quantidade incrível de informações genéticas acumuladas. Seis organismos-modelo, objetos de intenso estudo genético, são: *Drosophila melanogaster*, uma mosca-da-fruta, *Escherichia coli*, uma bactéria localizada no intestino dos humanos e de outros mamíferos, *Caenorhabditis elegans*, um verme nematódeo, *Arabidopsis thaliana*; *Mus musculus*, camundongo doméstico e *Saccharomyces cerevisiae*, levedura (**Figura 1.4**). Essas espécies são os organismos de escolha para muitos pesquisadores em genética, e seus genomas foram sequenciados como parte do Projeto Genoma Humano (ver *Capítulo 20*). Os ciclos de vida e as características genéticas desses organismos-modelo estão descritos com mais detalhes no Guia de Referência para Organismos-Modelo da Genética localizado no final deste livro (pp. 693-705). O Guia de Referência será uma fonte útil à medida que você encontrar esses organismos no livro.

À primeira vista, parece ser improvável que esse grupo de criaturas inferiores e, às vezes, desvalorizadas, seja candidato para organismos-modelo. Entretanto, todos apresentam ciclos de vida e características que os tornam particularmente adequados para estudo genético, incluindo um tempo de geração curto, mas com número exequível de descendentes, adaptabilidade ao ambiente laboratorial e a capacidade de serem armazenados e se reproduzirem a baixo custo. Outras espécies que são frequentemente objetos de pesquisa genética e são consideradas modelos genéticos incluem *Neurospora crassa* (mofo do pão), *Zea mays* (milho), *Danio rerio* (peixe-zebra) e *Xenopus laevis* (rã-de-unha africana). Embora não sejam considerados um modelo genético, os humanos também estão sujeitos a intenso exame genético minucioso; as técnicas especiais para análise genética dos humanos são discutidas no Capítulo 6.

A importância dos organismos-modelo da genética está ilustrada pelo uso do peixe-zebra para identificar os genes que afetam a pigmentação da pele nos seres humanos. Há muitos anos, os geneticistas sabiam que as diferenças na pigmentação entre os diferentes grupos étnicos humanos eram genéticas (**Figura 1.5 A**), mas os genes responsáveis por estas diferenças eram desconhecidos. O peixe-zebra se tornou um importante modelo nos estudos genéticos porque é um pequeno vertebrado que produz muitos filhotes e é fácil de criar em laboratório.

As células pigmentares da variante do peixe-zebra chamado de *dourado* são mais claras (**Figura 1.5 B**), porque as organelas que contêm o pigmento melanina, os melanossomos, estão em menor quantidade e tamanho e possuem menor concentração do pigmento.

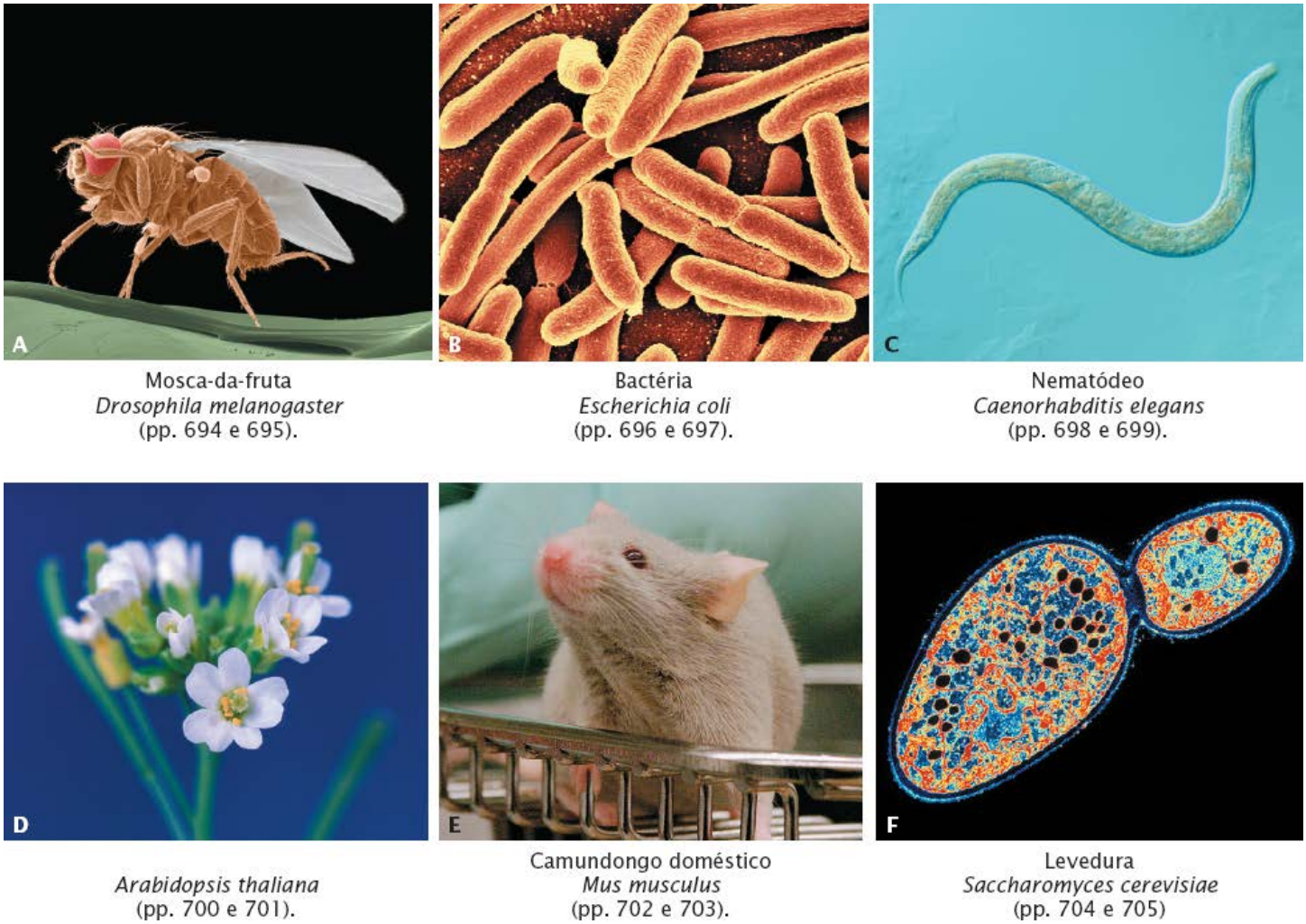


Figura 1.4 Os organismos-modelo da genética são espécies com características que os tornam úteis para a análise genética. (*Parte A*: SPL/Photo Researchers. *Parte B*: Pasieka/Photo Researchers, Inc. *Parte C*: Sinclair Stammers/Photo Researchers, Inc. *Parte D*: Peggy Greb/ARS USDA. *Parte E*: Joel Page/AP. *Parte F*: Biophoto Associates/Photo Researchers, Inc.)

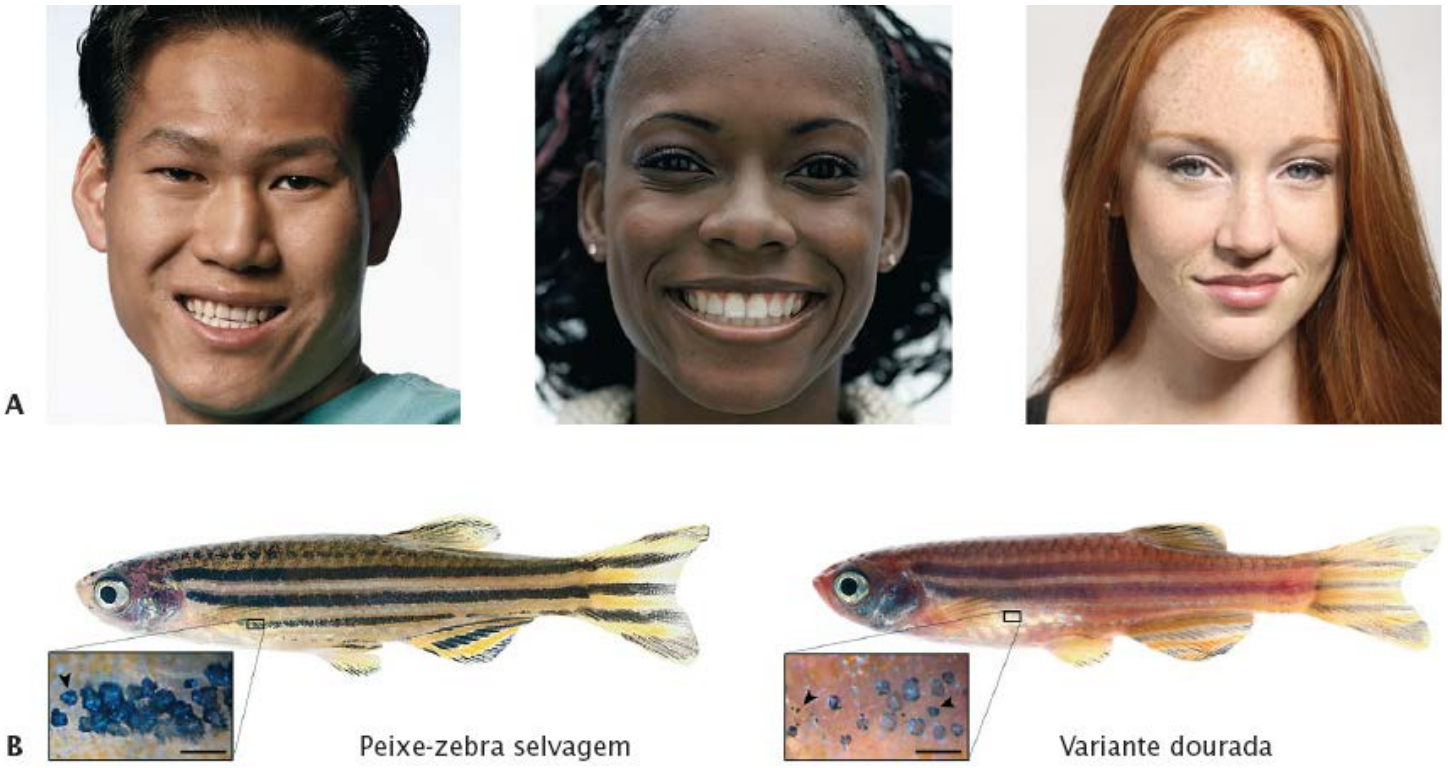


Figura 1.5 O peixe-zebra, um organismo-modelo da genética, é uma ferramenta usada para identificar os genes que codificam as diferenças de pigmentação entre os humanos. A. Os grupos étnicos humanos diferem no grau de pigmentação da pele. **B.** O fenótipo *dourado* do peixe-zebra é produzido por uma mutação no gene que controla quantidade, tamanho e concentração dos melanossomos, a organela do pigmento das células correspondentes aos melanócitos nos seres humanos. (*Parte A*: PhotoDisc/Getty

Images. *Parte B*: Keith Cheng/Jake Gittlen, Cancer Research Laboratories, Penn State Hershey College of Medicine.)

Keith Cheng *et al*. observaram que as mesmas alterações ocorriam nas pessoas com pele mais clara e desenvolveram uma hipótese na qual a pele clara nos seres humanos pode ser resultado de uma mutação que é semelhante à mutação *dourada* encontrada no peixe-zebra. Como é fácil manipular o peixe-zebra no laboratório, eles isolaram e sequenciaram o gene responsável pela mutação *dourada* e descobriram que ele codifica uma proteína que participa da captação de cálcio pelos melanossomos. Eles, então, pesquisaram em um banco de dados contendo todos os genes humanos conhecidos e descobriram um gene semelhante, o *SLC24A5*, com a mesma função nas células humanas. Quando os pesquisadores examinaram as populações humanas, eles descobriram que os europeus, de pele mais clara, possuíam uma forma desse gene, enquanto africanos, asiáticos do leste e nativos americanos, que têm pele mais escura, possuíam uma forma diferente do gene.

Outros genes também afetam a pigmentação nos humanos, como ilustrado pelas mutações no gene *OCA2*, que produzem albinismo entre os nativos americanos hopi. Contudo, parece que *SLC24A5* é responsável por 24 a 38% das diferenças de pigmentação entre africanos e europeus. Este exemplo ilustra a força dos organismos-modelo na pesquisa genética. Entretanto, não podemos esquecer que todos os organismos possuem características únicas, e, por vezes, a genética dos modelos não reflete completamente a genética de outros organismos.

Conceitos

As três principais divisões da genética são a genética da transmissão, genética molecular e genética populacional. A genética da transmissão examina os princípios da hereditariedade; a genética molecular lida com o gene e os processos celulares por meio dos quais as informações genéticas são transferidas e expressas; a genética populacional estuda a composição genética dos grupos de organismos e como a composição muda com a geografia e o tempo. Os organismos-modelo da genética são espécies que receberam ênfase especial na pesquisa genética; eles têm características que os tornam úteis para análise genética.

Checagem dos conceitos 2

O cavalo seria um bom organismo-modelo da genética? Sim ou não? Por quê?

1.2 Os humanos praticam genética há milhares de anos

Embora a ciência da genética seja jovem – cerca de 100 anos –, as pessoas usam os princípios genéticos há milhares de anos.

As primeiras técnicas genéticas e a compreensão da hereditariedade

A primeira evidência de que as pessoas compreendiam e aplicavam os princípios da hereditariedade é encontrada na domesticação de plantas e animais, que começou há aproximadamente 10.000 ou 12.000 anos no Oriente Médio. Os primeiros organismos domesticados incluem trigo, ervilhas, lentilhas, cevada, cães, bodes e carneiro (**Figura 1.6**). Sofisticadas *técnicas genéticas* já eram usadas no Oriente Médio há 4.000 anos. Os assírios e babilônios desenvolveram centenas de variedades de tamareiras que diferenciavam no tamanho, cor e sabor do fruto, e tempo de amadurecimento. Outras plantações e animais domesticados foram desenvolvidos por culturas na Ásia, África e Américas no mesmo período.

Figura 1.6 Os ancestrais praticavam técnicas genéticas na agricultura. O trigo moderno, com sementes maiores e em maior número, que não se espalham antes da colheita, foi produzido por meio do cruzamento de pelo menos três espécies selvagens diferentes. (Scott Bauer/ARS/USDA.)

As escrituras antigas demonstram que os primeiros humanos também tinham noção da sua própria hereditariedade. Os escritos hindus sagrados datados de 2.000 anos atrás atribuem muitas características ao pai e sugerem que diferenças entre os filhos são produzidas pela mãe. O Talmude, o livro judeu de leis religiosas com base nas tradições orais, que data de milhares de anos, apresenta uma extraordinária compreensão exata da herança da hemofilia. Ele afirma que, se uma mulher tiver dois filhos que morreram de sangramento após a circuncisão, qualquer outro filho que ela tenha, ou os filhos de suas filhas, não devem ser circuncidados. Esse aviso corresponde precisamente ao padrão ligado ao X da herança da hemofilia (discutido adiante, no *Capítulo 6*).

Os gregos antigos tinham cuidado com a reprodução e a hereditariedade humanas. Os filósofos gregos desenvolveram o conceito de **pangênese**, na qual partículas específicas, chamadas posteriormente de gêmulas, carregavam informações de várias partes do corpo para os órgãos reprodutores, as quais eram transmitidas para o embrião no momento da concepção (**Figura 1.7**). Embora incorreto, o conceito de pangênese teve muita influência e persistiu até o final dos anos 1800.

A pangênese levou os antigos gregos a proporem a noção da **herança das características adquiridas**, na qual as características adquiridas na vida de uma pessoa eram incorporadas nas informações hereditárias dela e transmitidas para seus filhos; por exemplo, pessoas que desenvolviam competência musical pelo estudo profundo teriam filhos que seriam naturalmente dotados com competência musical. A noção de herança das características adquiridas também não é mais aceita, mas permaneceu popular até o século 20.

Embora os romanos antigos tenham contribuído pouco para a compreensão da hereditariedade humana, eles desenvolveram várias técnicas para criação de animais e plantas; elas se baseavam mais na forma de tentativa e erro que em um conceito geral de hereditariedade. Poucas informações foram obtidas com a compreensão da genética nos 1.000 anos seguintes.

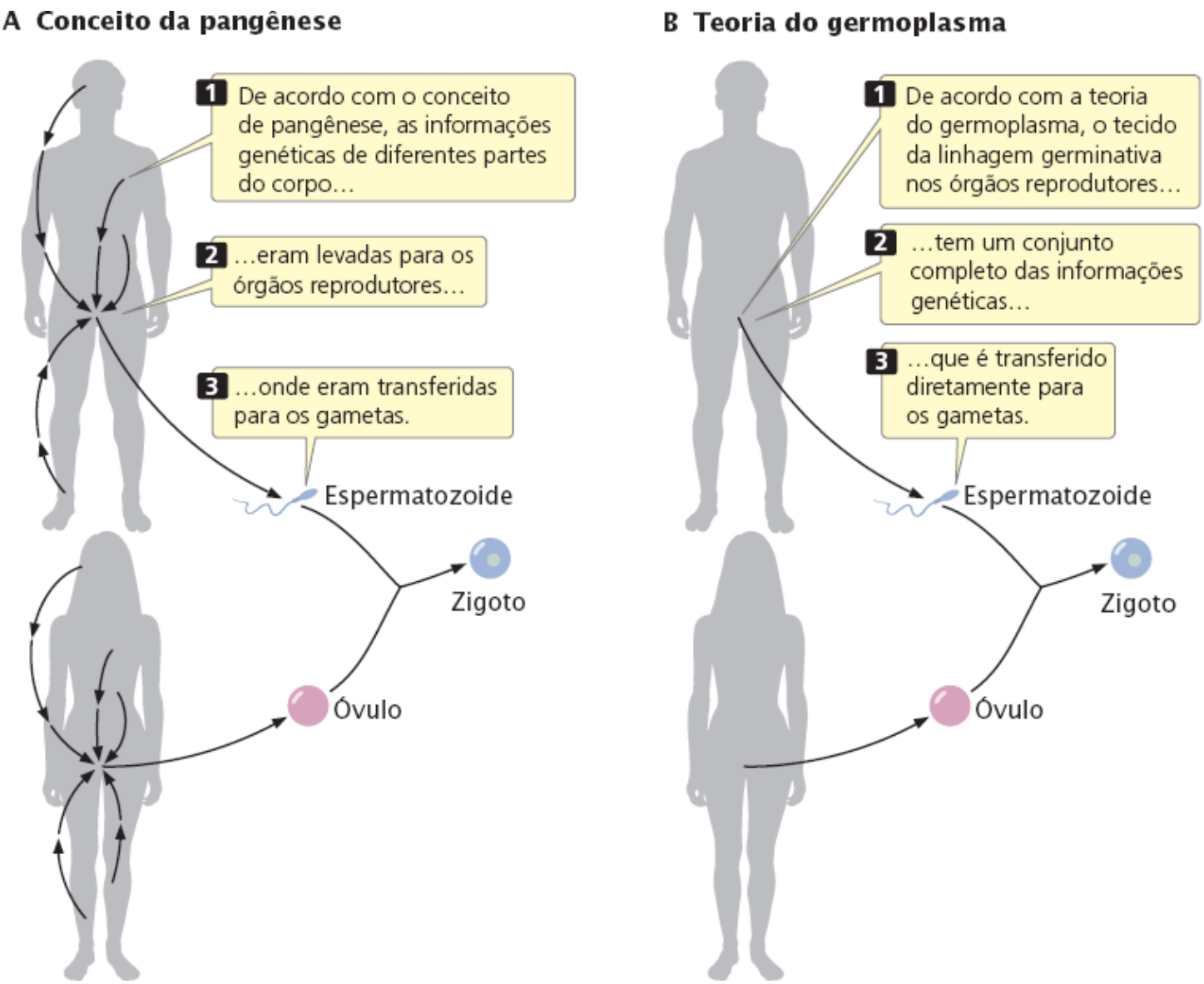


Figura 1.7 Pangênese, um conceito inicial de herança, comparado com a moderna teoria do germoplasma ou "plasma germinativo".

O desenvolvimento adicional da nossa compreensão sobre a hereditariedade ocorreu durante o século 17. Fabricantes de óculos holandeses criaram microscópios simples no final de 1500, permitindo que Robert Hooke (1635-1703) descobrisse as células em 1665. Os microscópios forneceram aos naturalistas uma nova e excitante perspectiva sobre a vida e, talvez, o entusiasmo excessivo com esse novo mundo minúsculo tenha dado origem à ideia do **pré-formacionismo**. De acordo com ele, dentro do óvulo ou espermatozoide existia um adulto em miniatura totalmente formado, um *homúnculo*, que simplesmente aumentava no curso do desenvolvimento. O pré-formacionismo significava que todas as características eram herdadas apenas do pai se o homúnculo estivesse no espermatozoide e da mãe se estivesse no óvulo. Embora surgissem muitas observações de que os descendentes têm uma mistura de características dos pais, o pré-formacionismo manteve-se como um conceito popular até os séculos 17 e 18.

Outra noção antiga de hereditariedade era a **herança por mistura**, que propunha que o descendente era uma combinação, ou mistura, das características dos pais. Esta ideia sugeriu que o próprio material genético se combinava, como os pigmentos azul e amarelo se misturavam para gerar a cor verde. Uma vez combinadas, as diferenças genéticas não poderiam ser separadas nas futuras gerações, assim como a cor verde não pode ser separada nos pigmentos azul e amarelo. Algumas características *realmente parecem ter* herança por mistura, entretanto, já sabemos que os genes individuais não se combinam.

O início da ciência genética

Em 1676, Nehemiah Grew (1641-1712) relatou que algumas plantas se reproduzem sexualmente ao usar o pólen das células sexuais masculinas. Com esta informação, vários botânicos começaram a fazer experimentos cruzando plantas e criando híbridos, incluindo Gregor Mendel (1822-1884, **Figura 1.8**) que descobriu os princípios básicos da hereditariedade. As conclusões de Mendel, que foram pouco divulgadas na comunidade científica por 35 anos, foram o alicerce da nossa

compreensão moderna da hereditariedade e, atualmente, ele é reconhecido como o pai da genética.

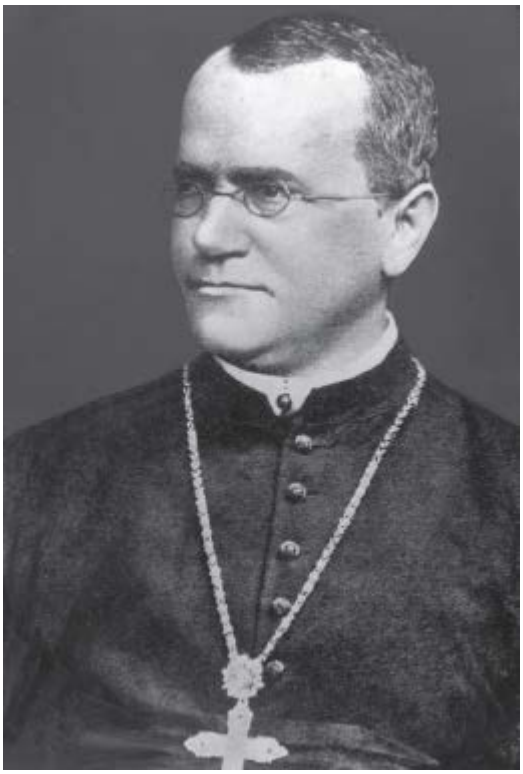


Figura 1.8 Gregor Mendel, o pai da genética moderna. Mendel foi o primeiro a descobrir os princípios de hereditariedade ao cruzar diferentes variedades de ervilhas e analisar a transmissão das características nas gerações subsequentes. (Hulton Archive/Getty Images.)

O desenvolvimento da citologia (o estudo das células) nos anos 1800 teve forte influência na genética. Robert Brown (1773-1858) descreveu o núcleo da célula em 1833. Junto com o trabalho de outros pesquisadores, Matthias Jacob Schleiden (1804-1881) e Theodor Schwann (1810-1882) propuseram o conceito de **teoria celular** em 1839. De acordo com ela, todas as formas de vida são compostas por células, que surgem apenas de outras preexistentes, e a célula é a unidade fundamental de estrutura e funcionamento dos organismos vivos. Os biólogos que tinham interesse na hereditariedade começaram a examinar as células para ver o que acontecia durante sua reprodução. Walther Flemming (1843-1905) observou a divisão dos cromossomos em 1879 e publicou uma excelente descrição da mitose. Em 1885, os biólogos reconheceram que o núcleo continha as informações da hereditariedade.

Charles Darwin (1809-1882), um dos biólogos mais influentes do século 19, desenvolveu a teoria da evolução por meio da seleção natural e publicou suas ideias no livro *A Origem das Espécies* em 1859. Darwin reconheceu que a hereditariedade era fundamental para a evolução e realizou amplos cruzamentos genéticos com pombos e outros organismos. Entretanto, ele nunca compreendeu a natureza da herança, e essa falta de compreensão foi uma omissão importante na sua teoria da evolução.

Na última metade do século 19, os citologistas demonstraram que o núcleo participava da fertilização. Perto da virada do século, August Weismann (1834-1914) finalmente derrubou a noção de herança das características adquiridas. Ele cortou os rabos de camundongos por 22 gerações consecutivas e mostrou que o rabo nos descendentes permanecia longo. Weismann propôs a **teoria do germoplasma** (“plasma germinativo”), que defende que as células nos órgãos reprodutores carregam um conjunto completo de informações genéticas que é passado para o óvulo e o espermatozoide (ver Figura 1.7 B).

O ano de 1900 foi um divisor de águas na história da genética. A publicação essencial de Gregor Mendel, de 1866, sobre suas experiências com ervilhas, que revelou os princípios da hereditariedade, foi redescoberta, como estudado em mais detalhes no Capítulo 3. A importância das suas conclusões foi identificada e outros biólogos imediatamente começaram a realizar estudos genéticos semelhantes em camundongos, galinhas e outros organismos. Os resultados dessas investigações mostraram que, de fato, muitas características seguiam as regras de Mendel. Alguns dos conceitos iniciais de hereditariedade estão resumidos no **Quadro 1.1**.

Após a aceitação da teoria da hereditariedade de Mendel, em 1902, Walter Sutton (1877-1916) propôs que os genes, as unidades da herança, estavam localizados nos cromossomos. Thomas Hunt Morgan (1866-1945) descobriu o primeiro mutante genético da mosca-da-fruta em 1910 e usou o inseto para desenvolver muitos detalhes da genética da transmissão. Ronald A. Fisher (1890-1962), John B. S. Haldane (1892-1964) e Sewall Wright (1889-1988) criaram os alicerces para a genética populacional nos anos de 1930 ao integrar a genética mendeliana e a teoria da evolução.

Quadro 1.1 Conceitos iniciais da hereditariedade.

Conceito	Proposto	Certo ou errado
Pangênese	As informações genéticas migram por diferentes partes do corpo para os órgãos reprodutores	Errado
Herança das características adquiridas	Características adquiridas são incorporadas em informações hereditárias	Errado
Pré-formacionismo	Organismo em miniatura reside nas células sexuais e todas as características são herdadas do pai ou da mãe	Errado
Herança por mistura	Os genes se combinam e misturam	Errado
Teoria do germoplasma	Todas as células contêm um conjunto completo de informações genéticas	Correto
Teoria celular	Toda forma de vida é composta por células e as células surgem apenas a partir de outras células	Correto
Herança mendeliana	As características são herdadas de acordo com princípios definidos	Correto

Os geneticistas começaram a usar bactérias e vírus nos anos 1940, e a rápida reprodução e os sistemas genéticos simples dos organismos permitiram o estudo detalhado da organização e estrutura dos genes. Ao mesmo tempo, surgiam evidências de que o DNA era o depósito de informações genéticas. James Watson (n.* 1928) e Francis Crick (1916-2004), junto com Maurice Wilkins (1916-2004) e Rosalind Franklin (1920-1958), descreveram a estrutura tridimensional do DNA em 1953 e conduziram à era da genética molecular.

Em 1966, a estrutura química do DNA e o sistema por meio do qual ele determina a sequência de aminoácidos das proteínas foram revelados. Os avanços na genética molecular levaram aos primeiros experimentos de DNA recombinante em 1973, que liderou outra revolução na pesquisa genética. Walter Gilbert (n. 1932) e Frederick Sanger (n. 1918) desenvolveram métodos para sequenciar o DNA em 1977. A reação em cadeia da polimerase, uma técnica para amplificar rapidamente mínimas quantidades de DNA, foi desenvolvida por Kary Mullis (n. 1944) e outros em 1983. Em 1990, a terapia genética foi usada pela primeira vez para tratar doença genética, nos EUA, e o Projeto Genoma Humano foi iniciado. Em 1995, foi determinada a primeira sequência completa de DNA de um organismo de vida livre (*free-living*) – a bactéria *Haemophilus influenzae* – e a primeira sequência completa de um organismo eucariótico (levedura) foi reportada um ano depois. Um esboço grosseiro do genoma humano foi descrito em 2000 (ver *Capítulo 20*) com a sequência essencialmente completa em 2003, abrindo uma nova era na genética (**Figura 1.9**). Hoje, os genomas de vários organismos estão sendo sequenciados, analisados e comparados.

Resolva os Problemas 22 e 23

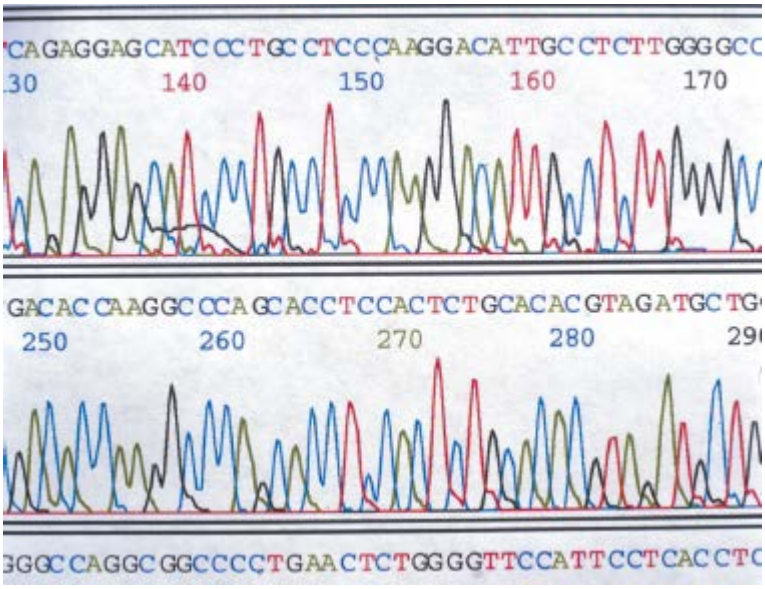


Figura 1.9 Em 2003, o genoma humano foi completamente sequenciado. Um cromatograma de uma pequena porção do genoma humano. (Science Museum/SSPL.)

O futuro da genética

Atualmente, a genética tem vários avanços acontecendo e permanece na vanguarda da pesquisa biológica. São usados métodos novos e rápidos de sequenciamento de DNA para sequenciar os genomas de várias espécies, desde morangos a borboletas e elefantes. Recentemente, esses métodos foram usados para reconstruir o genoma inteiro de um feto abortado a partir do DNA fetal circulante no sangue da mãe, fornecendo o potencial para teste genético pré-natal não invasivo. A análise de DNA de ossos antigos demonstra que várias espécies diferentes de humanos perambularam pela Terra há 30.000 anos. As técnicas genéticas modernas e poderosas são usadas para identificar genes que influenciam características importantes na agricultura como tamanho do gado, domesticação nas galinhas, velocidade dos cavalos de corrida e formato da folha do milho. A análise de DNA é usada atualmente para identificar e condenar criminosos ou provar a inocência de suspeitos.

O poder dos novos métodos para identificar e analisar os genes é ilustrado por meio dos estudos genéticos do infarto do miocárdio nos humanos. Os médicos reconhecem que existe um padrão familiar nos infartos do miocádio que afetam famílias, mas encontrar genes específicos que contribuem para um risco maior desse infarto cardíaco tem sido, até o momento, uma tarefa difícil. Em 2009, uma equipe internacional de geneticistas examinou o DNA de 26.000 pessoas em 10 países para diferenças únicas de nucleotídio no DNA (chamados de polimorfismos de único nucleotídio ou SNPs [do inglês, *single nucleotide polymorphisms*]) que podiam estar associados a um risco maior de infarto do miocárdio. Este e outros estudos semelhantes identificaram vários novos genes que afetam o risco de coronariopatia e infartos do miocárdio precoces. Estas descobertas permitem identificar as pessoas predispostas a um infarto, o que permite uma intervenção precoce que pode evitá-lo. As análises de SNPs estão ajudando a localizar os genes que afetam todos os tipos de características, desde a cor dos olhos e peso a glaucoma e câncer.

As informações sobre diferenças na sequência entre os organismos também é a fonte de novas percepções sobre a evolução. Por exemplo, recentemente os cientistas analisaram a sequência de DNA de 26 genes para construir uma árvore evolutiva abrangente dos mamíferos. A árvore revela muitas características interessantes da evolução dos mamíferos. Uma delas é que os mamíferos marinhos (baleias e golfinhos) são mais próximos dos hipopótamos.

Recentemente, os cientistas descobriram que alterações ao DNA e à estrutura do cromossomo que não envolvem a sequência de bases do DNA têm um papel importante na expressão do gene. Essas alterações, chamadas de alterações epigenéticas, afetam nossa aparência, comportamento e saúde, sendo, atualmente, foco de intensa pesquisa. Outros estudos demonstram que o RNA é crucial em muitos aspectos da função dos genes. A descoberta, no final de 1990, de pequenas moléculas de RNA chamadas de pequenos RNA de interferência e micro-RNAs levou ao reconhecimento de que estas moléculas têm papéis centrais na expressão gênica e desenvolvimento. Novos *microchips* genéticos que analisam simultaneamente milhares de moléculas de RNA estão fornecendo informações sobre as atividades de milhares de genes em uma célula, permitindo uma imagem detalhada de como as células respondem aos sinais externos, estresses ambientais e doenças como câncer. No campo da proteômica, poderosos programas computacionais estão sendo usados para modelar a estrutura e função de proteínas a partir das informações da sequência de DNA. Todas essas informações possibilitam melhor compreensão dos numerosos processos biológicos e relações evolutivas. Uma enchente de novas informações genéticas requer o desenvolvimento contínuo de sofisticados programas computacionais para armazenar, restaurar, comparar e analisar os dados genéticos, processo que deu origem ao campo da bioinformática, uma união de biologia molecular e ciência computacional.

Com um custo mais acessível para o sequenciamento, o foco do sequenciamento do DNA se desloca para os genomas de diferentes espécies para diferenças individuais dentro da espécie. Em um futuro não muito distante, é provável que cada pessoa tenha uma cópia completa de sua sequência de genoma, que pode ser usada para auxiliar a avaliar o risco de adquirir diferentes doenças e traçar o tratamento se elas surgirem. O uso da genética na agricultura continuará a melhorar a produtividade das culturas e animais domésticos, ajudando a alimentar a população mundial no futuro. Este alcance abrangente da genética traz questões éticas, sociais e econômicas importantes.

Esse breve panorama da história da genética não pretende ser abrangente; é planejado para fornecer uma noção do passo acelerado dos avanços na genética. Aprenderemos nos próximos capítulos mais sobre os experimentos e os cientistas que ajudaram a moldar a disciplina da genética.

Conceitos

Os humanos aplicaram a genética pela primeira vez na domesticação de plantas e animais entre 10.000 e 12.000 anos atrás. O desenvolvimento na hibridização das plantas e citologia nos séculos 18 e 19 formaram a base para o campo da genética hoje. Após o trabalho de Mendel ser redescoberto em 1900, a ciência da genética se desenvolveu rapidamente e hoje é uma das áreas mais ativas da ciência.

Checagem dos conceitos 3

Como os desenvolvimentos na citologia no século 19 contribuíram para nossa compreensão moderna da genética?

1.3 Alguns conceitos são fundamentais para o início de nossa jornada na Genética

Sem dúvida, você aprendeu alguns princípios de genética em outras aulas de Biologia. Vamos revisar alguns desses conceitos fundamentais.

Existem dois tipos básicos de células | Eucariótica e procariótica. Estruturalmente, as células consistem em dois tipos básicos, embora do ponto de vista da evolução, a história seja mais complexa (ver *Capítulo 2*). As células procarióticas não têm membrana celular e, em geral, não têm organelas celulares ligadas a membranas, enquanto as células eucarióticas são mais complexas, têm um núcleo e organelas ligadas a membranas, como cloroplastos e mitocôndrias.

O gene é a unidade fundamental da hereditariedade. A forma precisa na qual um gene é definido varia, dependendo do contexto biológico. No nível mais simples, podemos pensar no gene como uma unidade de informações que codifica uma característica genética. Expandimos esta definição à medida que aprendemos mais sobre o que são os genes e como eles funcionam.

Os genes se apresentam como formas múltiplas chamadas de alelos. Um gene que determina uma característica pode existir em várias formas. Por exemplo, um gene para cor de pelo nos gatos pode existir como um alelo que codifica o pelo preto ou um alelo que codifica o pelo laranja.

Os genes conferem os fenótipos. Um dos conceitos mais importantes na genética é a diferença entre características e genes. As características não são herdadas diretamente. Ao contrário, os genes são herdados e, junto com os fatores ambientais, determinam a expressão das características. As informações genéticas que um organismo individual tem são seu genótipo, a característica observável é seu fenótipo. Por exemplo, o albinismo observado em alguns hopis é um fenótipo e as informações nos genes *OCA2* que causam o albinismo são o genótipo.

As informações genéticas residem no DNA e no RNA. As informações genéticas são codificadas na estrutura molecular dos ácidos nucleicos, que são de dois tipos: ácido desoxirribonucleico (DNA) e ácido ribonucleico (RNA). Os ácidos nucleicos são polímeros de unidades repetidas chamadas de nucleotídios, e cada um deles consiste em um açúcar, um fosfato e uma base nitrogenada. Existem quatro tipos de bases nitrogenadas no DNA: adenina (A), citosina (C), guanina (G) e timina (T). A sequência dessas bases codifica as informações genéticas. O DNA consiste em duas fitas de nucleotídios complementares. A maior parte dos organismos carrega suas informações genéticas no DNA, mas alguns vírus carregam essas informações no RNA. As quatro bases nitrogenadas do RNA são adenina, citosina, guanina e uracila (U).

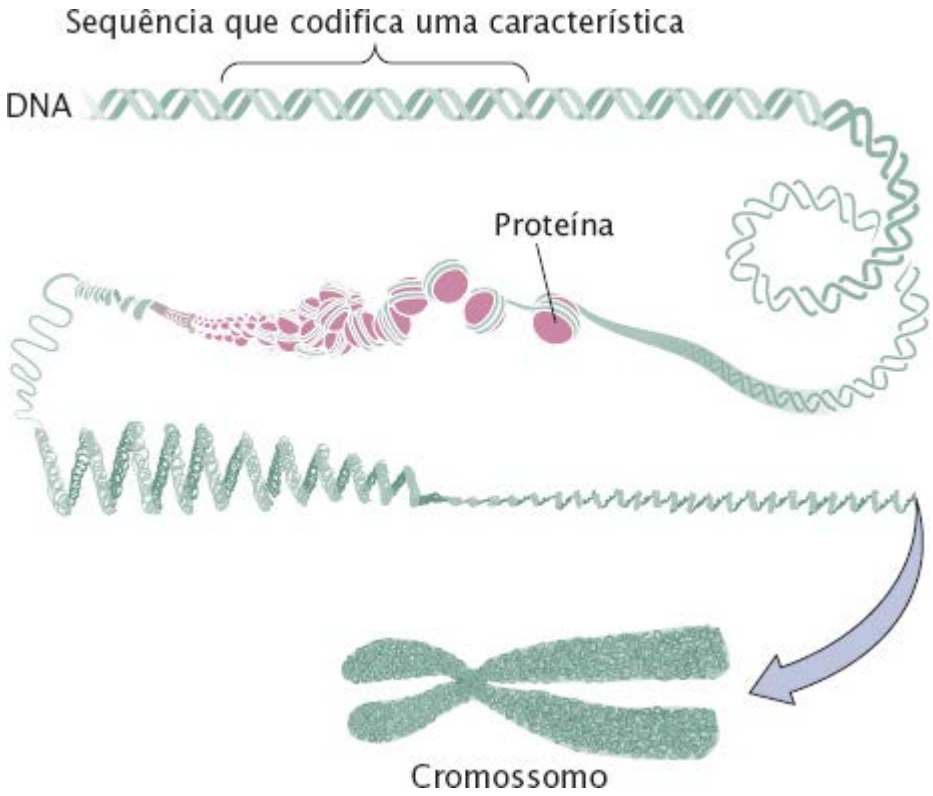


Figura 1.10 **Os genes residem nos cromossomos.**

Os genes estão localizados nos cromossomos. Os veículos das informações genéticas dentro de uma célula são os cromossomos (**Figura 1.10**), que consistem em DNA e proteína associados. As células de cada espécie têm um número característico de cromossomos; por exemplo, normalmente as células bacterianas têm um único cromossomo, as células humanas têm 46 e as células do pombo, 80. Cada cromossomo carrega vários genes.

Os cromossomos são separados por meio de processos de mitose e meiose. Esses processos garantem que exista um conjunto completo de cromossomos de um organismo em cada célula oriunda da divisão celular. A mitose é a separação dos cromossomos na divisão das células somáticas (não sexuais). A meiose é o pareamento e a separação dos cromossomos na divisão das células sexuais para produzir os gametas (células reprodutoras).

As informações genéticas são transferidas do DNA para o RNA e para as proteínas. Muitos genes codificam características ao descrever a estrutura das proteínas. As informações genéticas são transcritas primeiro do DNA para o RNA e então o RNA é traduzido na sequência de aminoácidos de uma proteína.

As mutações são alterações permanentes nas informações genéticas, que podem ser transmitidas de célula para célula ou do pai para os descendentes. As mutações genéticas afetam as informações genéticas de apenas um único gene. Já as mutações dos cromossomos alteram o número ou a estrutura dos cromossomos e, portanto, em geral, afetam muitos genes.

Algumas características são afetadas por múltiplos fatores. Algumas características são afetadas por múltiplos genes que interagem de formas complexas com os fatores ambientais. A altura do ser humano, por exemplo, é afetada por muitos genes, assim como fatores ambientais como nutrição.

A evolução é uma mudança genética. A evolução pode ser vista como um processo de duas etapas: primeiro, a variação genética surge e, depois, algumas variantes genéticas aumentam em frequência, enquanto outras diminuem. **Resolva o Problema 24**.

Resumo dos conceitos

- A genética é central para a vida de todos: ela influencia as características físicas, a personalidade, inteligência e suscetibilidade da pessoa a inúmeras doenças
- Ela é importante na agricultura, na indústria farmacêutica e na Medicina. É central para o estudo da Biologia
- Todos os organismos usam sistemas genéticos semelhantes. A variação genética é o alicerce da evolução e é crítico para

compreender todas as formas de vida

- O estudo da genética pode ser, de modo geral, dividido em genética da transmissão, genética molecular e genética populacional
- Os organismos-modelo da genética são espécies sobre as quais existem muitas informações genéticas porque eles têm características que os tornam particularmente receptivos à análise genética
- O uso da genética pelos humanos começou com a domesticação de plantas e animais
- Os gregos antigos desenvolveram o conceito de pangênese e a herança de características adquiridas; no entanto, ambos foram refutados posteriormente. Os romanos antigos desenvolveram medidas práticas para a procriação de plantas e animais
- O pré-formacionismo sugeria que uma pessoa herdava todas as características do pai ou da mãe. A herança combinada propôs que o descendente tem uma mistura das características dos pais. Posteriormente, foi demonstrado que estas ideias estavam erradas
- Ao estudar os descendentes de cruzamentos entre variedades de ervilhas, Gregor Mendel descobriu os princípios da hereditariedade. O desenvolvimento da citologia no século 19 levou à compreensão de que o núcleo da célula é o local da hereditariedade
- Em 1900, os princípios de Mendel foram redescobertos. A genética populacional foi estabelecida no início dos anos 1930, seguida de perto pela genética bioquímica e genética bacteriana e viral. A estrutura do DNA foi descoberta em 1953, estimulando o surgimento da genética molecular
- Existem dois tipos básicos de células: procarióticas e eucarióticas
- Os genes que determinam uma característica são chamados de genótipo, a característica que eles produzem é o fenótipo
- Os genes estão localizados nos cromossomos, que são constituídos por ácidos nucleicos e proteínas e são divididos nas células-filhas por meio do processo de mitose ou meiose
- As informações genéticas são expressas pela transferência de informações do DNA para RNA e para as proteínas
- A evolução requer uma mudança genética nas populações.

Termos importantes

Genética da transmissão
Genética molecular
Genética populacional
Genoma
Herança das características adquiridas
Herança por mistura
Organismo-modelo da genética
Pangênese
Pré-formacionismo
Teoria celular
Teoria do germoplasma

Respostas da Checagem dos conceitos

1. d.
2. Não, porque é caro abrigar, alimentar e reproduzir cavalos, eles têm poucos filhotes e o tempo da sua geração é muito longo.
3. O desenvolvimento da citologia nos anos 1800 levou à identificação de partes da célula, incluindo o seu núcleo e os cromossomos. A teoria celular se concentra na atenção dos biólogos sobre a célula, eventualmente levando à conclusão de que o núcleo tem as informações da hereditariedade.

Questões de compreensão

As respostas às questões precedidas por um asterisco podem ser encontradas no final do livro.

Seção 1.1

*1. Como a cultura hopi contribuiu para a alta incidência de albinismo entre os membros dessa tribo?
2. Destaque algumas formas nas quais a genética é importante para todos.
3. Dê pelo menos três exemplos do papel da genética na sociedade hoje em dia.
4. Explique resumidamente porque a genética é crucial para a Biologia moderna.
5. Liste as três subdisciplinas tradicionais da genética e resuma o que cada uma estuda.

6. Quais são as características dos organismos-modelo da genética que os tornam úteis para os estudos de genética?

Seção 1.2

7. Quando e onde a agricultura surgiu pela primeira vez? Qual o papel da genética no desenvolvimento dos primeiros animais e plantas domesticados?
8. Descreva a noção de pangênese e explique sua diferença da teoria do germoplasma.
9. O que o conceito de herança das características adquiridas propõe e como ele está relacionado com a noção de pangênese?
10. O que é pré-formacionismo? O que ele fala sobre como as características são herdadas?
11. Defina a herança por mistura e a diferencie do pré-formacionismo.
12. Como o desenvolvimento na botânica nos séculos 17 e 18 contribuiu para o surgimento da genética moderna?
13. Liste alguns avanços na genética feitos no século 20.
14. Explique resumidamente a contribuição de cada uma das pessoas a seguir para o estudo da genética.
    a. Matthias Schleiden e Theodor Schwann.
    b. August Weismann.
    c. Gregor Mendel.
    d. James Watson e Francis Crick.
    e. Kary Mullis.

Seção 1.3

15. Quais são os dois tipos básicos de células (da perspectiva estrutural) e quais são suas diferenças?
16. Descreva as relações entre genes, DNA e cromossomos.

Questões e problemas aplicados

Seção 1.1

*17. Qual a relação entre genética e evolução?
*18. Para cada um dos tópicos de genética, indique se ele se concentra na genética da transmissão, genética molecular ou genética populacional.
    a. A análise de linhagens para determinar a probabilidade de alguém herdar uma característica.
    b. Estudo de pessoas em uma pequena ilha para determinar por que uma forma genética de asma é prevalente na ilha.
    c. Efeito do cruzamento não aleatório sobre a distribuição dos genótipos entre um grupo de animais.
    d. O exame das sequências de nucleotídio encontradas nas terminações dos cromossomos.
    e. Os mecanismos que garantem um alto grau de precisão na replicação do DNA.
    f. Estudo de como a herança das características codificada pelos genes nos cromossomos sexuais (características ligadas ao sexo) difere da herança das características codificadas por genes em cromossomos não sexuais (características autossômicas).
19. Descreva algumas das formas pelas quais a sua própria composição genética o afeta como pessoa. Seja o mais específico que puder.
20. Descreva pelo menos uma característica que parece acompanhar sua família (aparece em múltiplos membros da família). Esta característica acompanha sua família porque é uma característica herdada ou é causada por fatores ambientais que são comuns aos membros da família? Como você poderia diferenciar entre estas possibilidades?

Seção 1.2

*21. A genética é ao mesmo tempo uma ciência muito antiga e muito nova. Explique o que significa esta afirmação.
*22. Combine a descrição (de *a* a *d*) com a teoria ou conceito correto listado a seguir.
    Pré-formacionismo.
    Pangênese.
    Teoria do germoplasma.
    Herança das características adquiridas.
    a. Cada célula reprodutora tem um conjunto completo de informações genéticas.
    b. Todos as características são herdadas do pai ou da mãe.
    c. Todas as informações genéticas podem ser alteradas pelo uso de uma característica.
    d. Diferentes tipos de células têm diferentes informações genéticas.
*23. Compare e contraste as seguintes ideias sobre herança.

a. Pangênese e teoria do germoplasma.
b. Pré-formacionismo e herança por mistura.
c. Herança das características adquiridas e nossa teoria moderna de hereditariedade.

Seção 1.3

*24. Compare e contraste os itens a seguir:
a. Células eucarióticas e procarióticas.
b. Gene e alelo.
c. Genótipo e fenótipo.
d. DNA e RNA.
e. DNA e cromossomo.

Questões desafiadoras

Introdução

*25. É provável que o tipo de albinismo surgido com alta frequência entre os hopi seja o albinismo oculocutâneo do tipo II, causado por um defeito no gene *OCA2* no cromossomo 15. Pesquise na internet para determinar como o fenótipo deste tipo de albinismo difere dos fenótipos de outras formas de albinismo nos humanos e os genes mutantes que resultam nestes fenótipos. Dica: visite o *site* Online Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/omim/) e pesquise o banco de dados sobre albinismo.

Seção 1.1

26. Hoje já temos muito conhecimento sobre a genética dos humanos e eles são o foco de muitos estudos de genética. Quais motivos explicam por que os humanos são o foco de intenso estudo genético?

Seção 1.3

*27. Suponha que exista vida em outro lugar no universo. Todas as formas de vida têm algum tipo de informação genética, mas os genomas alienígenas podem não consistir em ácidos nucleicos e ter algumas características como as encontradas nos genomas nas formas de vida na Terra. Quais seriam as características comuns de todos os genomas, não importa onde existam?
28. Escolha uma questão ética ou social entre as letras *a* até *d* e de sua opinião sobre o assunto. Para obter informações do histórico, você pode ler artigos sobre ética, disponibilizados em diversos sites.
a. A composição genética de uma pessoa deve ser usada para determinar sua elegibilidade para seguro de vida?
b. As empresas de biotecnologia devem ser capazes de patentear genes recém-sequenciados?
c. A terapia gênica deve ser usada em pessoas?
d. O teste genético deve ser disponibilizado para pesquisar distúrbios hereditários, para os quais não existe tratamento ou cura?
29. Uma mulher de 45 anos é submetida ao teste genético e descobre que corre alto risco de desenvolver câncer de cólon e doença de Alzheimer. Como os filhos têm 50% dos seus genes, eles também correm maior risco para estas doenças. Ela tem a obrigação moral ou legal de contar a seus filhos e outros parentes próximos sobre os resultados do seu teste genético?
30. Suponha que você seja submetido a um teste genético aos 18 anos quanto à suscetibilidade de uma doença genética que não apareceria antes da meia-idade e para a qual não há tratamento disponível.
a. Quais seriam os possíveis motivos para fazer esse teste e quais seriam os possíveis motivos para não fazê-lo?
b. Você gostaria de fazer este teste? Explique seu motivo.

*Para n., lê-se nascido em.

2 Cromossomos e Reprodução Celular

O enigma dos cegos

Em uma charada muito conhecida, dois cegos entram por acaso em uma loja de departamentos ao mesmo tempo, vão ao mesmo balcão e pedem cinco pares de meias, cada par de uma cor diferente. O vendedor fica tão surpreso com essa estranha coincidência que coloca todos os dez pares (dois de cor preta, dois de cor azul, dois de cor cinza, dois de cor marrom e dois de cor verde) em uma única bolsa de compras, entrega uma bolsa com as meias para um dos cegos e uma bolsa vazia para o outro. No entanto, os dois cegos se encontram na rua, onde descobrem que uma das bolsas tinha todos os dez pares de meias. Como os cegos, sem ver e sem receber ajuda de outra pessoa, separam as meias de modo que cada um vá para casa com exatamente cinco pares de meias de diferentes cores? Você pode solucionar esse enigma?

Por uma interessante coincidência, as células enfrentam o mesmo dilema que os cegos dessa charada. A maior parte dos organismos tem dois conjuntos de informações genéticas, um herdado do pai e outro da mãe. Antes da divisão celular, o DNA em cada cromossomo é replicado; após a replicação, existem duas cópias, chamadas de cromátides-irmãs, de cada cromossomo. No final da divisão celular, é crítico que cada uma das duas novas células receba uma cópia completa do material genético, assim como cada cego precisa ir para casa com um conjunto completo de meias.

A solução do enigma é simples. As meias são vendidas em pares; as duas meias de um par estão presas por um fio. À medida que um par é retirado da bolsa, cada homem segura uma meia diferente do par e puxa no sentido oposto. Quando as meias são puxadas juntas, um dos homens pode pegar uma faca pequena e cortar o fio que prende o par. Cada homem coloca sua meia separada na própria bolsa. No final do processo, cada bolsa terá exatamente duas meias pretas, duas azuis, duas cinza, duas marrons e duas verdes.*

As células usam uma solução semelhante para separar seus cromossomos nas células-filhas. Conforme aprenderemos neste capítulo, os cromossomos replicados se alinham no centro de uma célula em divisão e, como as meias na charada, as cromátides-irmãs de cada cromossomo são puxadas em direções opostas. Como a linha que une as duas meias em um par, uma molécula chamada coesina mantém as cromátides-irmãs juntas até serem separadas por uma faca molecular chamada de separase. Os dois cromossomos resultantes se separam, e a célula se divide, garantindo que um conjunto completo de cromossomos seja depositado em cada célula.

Nessa analogia, os cegos e as células diferem em uma questão crítica: se os cegos cometerem um erro, um cego fica com uma meia a mais, e o outro, com uma a menos, mas sem grande prejuízo. O mesmo não se pode dizer sobre as células humanas. Erros na separação do cromossomo, que produzem células com mais ou menos cromossomos, são, em muitos casos, catastróficos, resultando em câncer, aborto ou – em alguns casos – uma criança com graves limitações.

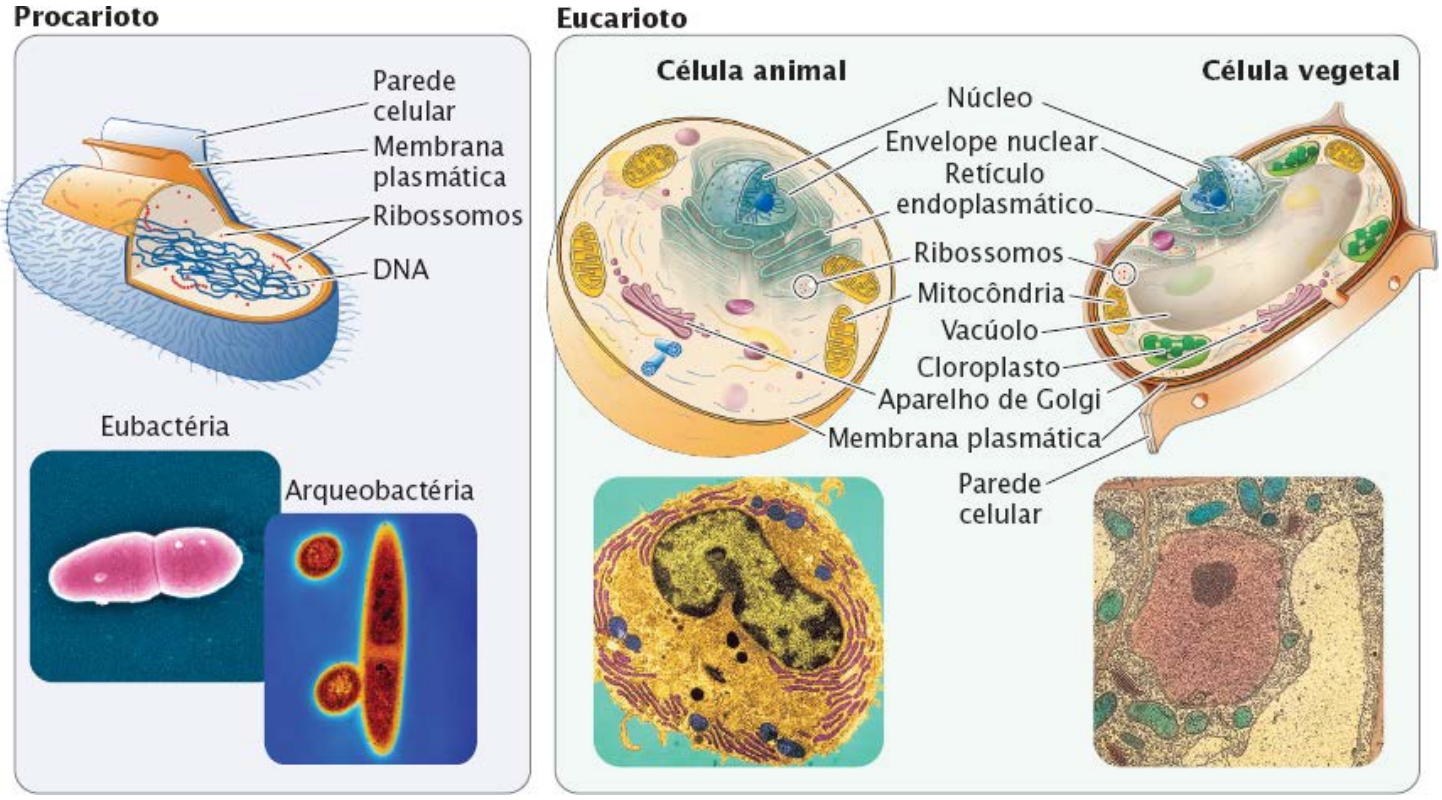
Este capítulo explora o processo de reprodução celular e como um conjunto completo de informações genéticas é transmitido para novas células. Nas células procarióticas, a reprodução é relativamente simples porque elas têm um único cromossomo. Nas eucarióticas, os múltiplos cromossomos têm de ser copiados e distribuídos para cada uma das novas células, tornando a reprodução celular mais complexa. A divisão celular nos eucariotos ocorre por meio de mitose e meiose, processos que são o alicerce para boa parte da genética.

Entender os processos de mitose e meiose requer mais que apenas memorizar as sequências de eventos que ocorrem em cada estágio, embora eles sejam importantes. A chave é compreender como as informações genéticas são distribuídas no curso da reprodução celular por meio de interação dinâmica de síntese de DNA, movimento dos cromossomos e divisão celular. Esses processos levam à transmissão das informações genéticas e são a base das semelhanças e diferenças entre os genitores e sua prole.

2.1 As células procarióticas e eucarióticas apresentam diferenças em

várias características genéticas

Tradicionalmente, os biólogos classificam todos os organismos vivos em dois grupos principais, *procariotos* e *eucariotos* (**Figura 2.1**). Um **procarioto** é um organismo unicelular com estrutura celular relativamente simples. Um **eucarioto** tem uma estrutura celular compartimentada, com os componentes envoltos por membranas intercelulares, sendo os membros dessa categoria unicelulares ou multicelulares.



	Células procarióticas	Céulas eucarióticas
Núcleo	Ausente	Presente
Diâmetro da célula	Relativamente pequenas, de 1 a 10 µm	Relativamente grande, de 10 a 100 µm
Genoma	Geralmente uma molécula de DNA circular	Múltiplas moléculas de DNA lineares
DNA	Não complexado com histonas nas eubactérias; algumas histonas na archaea	Complexado com histonas
Quantidade de DNA	Relativamente pequena	Relativamente grande
Organelas envoltas por membrana	Ausentes	Presentes

Figura 2.1 As células procarióticas e eucarióticas têm estruturas diferentes. (Fotografias [*da esquerda para a direita*] Dr. Gary D. Gaugler/Newscom; Dr. Kari Lounatmaa/Science Source; W. Baumeister/Science Photo Library/Photo Researchers; G. Murti/Phototake; Biophoto Associates/Photo Researchers.)

Pesquisas indicam que uma divisão das formas de vida em dois grupos principais não é tão simples. Embora apresentem semelhanças na estrutura celular, os procariotos incluem pelo menos dois tipos fundamentalmente diferentes de bactérias, as **eubactérias** (bactérias verdadeiras) e as **archaea** (bactérias antigas). Um exame das sequências de DNA equivalentes revela que as eubactérias e as archaea estão tão distantes uma das outras quanto dos eucariotos. Embora as eubactérias e archaea tenham estrutura celular semelhante, alguns processos genéticos na archaea (como a transcrição) são mais semelhantes aos dos eucariotos, estando até mais próximas do ponto de vista evolutivo dos eucariotos que das eubactérias. Assim, do ponto de vista evolutivo, existem três grupos principais de organismos: eubactérias, archaea e eucariotos. A distinção entre procarioto e eucarioto será comum neste livro, mas importantes diferenças entre eubactérias e archaea também serão apresentadas.

Da perspectiva da genética, uma diferença importante entre as células procarióticas e eucarióticas é que um eucarioto tem um envelope celular, que circunda o material genético para formar um **núcleo** e separa o DNA do restante do material celular. Nas células procarióticas, o material genético mantém contato próximo com os outros componentes da célula – uma propriedade que tem importantes consequências na forma como os genes são controlados.

Outra diferença fundamental entre os procariotos e eucariotos é no acondicionamento do seu DNA. Nos eucariotos, o DNA está intimamente associado a uma classe especial de proteínas, as **histonas**, para formar cromossomos bem acondicionados (**Figura 2.2**). Este complexo de DNA e histonas é chamado de **cromatina**, que é a composição dos cromossomos eucarióticos. As proteínas histona limitam a acessibilidade das enzimas e outras proteínas que copiam e leem o DNA, mas permitem que o DNA se "encaixe" no núcleo. O DNA eucariótico tem de ser separado das histonas antes que as informações genéticas do DNA possam ser acessadas. Archaea também têm algumas histonas que se complexam com o DNA, mas a estrutura de sua cromatina é diferente da encontrada nos eucariotos. As eubactérias não têm histonas; desse modo, seu DNA não existe em uma disposição bem organizada, bem acondicionada como a encontrada nas células eucarióticas. Os processos de cópia e leitura do DNA são mais simples nas eubactérias.

Os genes das células procarióticas são, geralmente, formados por uma molécula única, circular, de DNA – o cromossomo de uma célula procariótica. Nas células eucarióticas, os genes estão localizados em múltiplas moléculas, em geral lineares, de DNA (múltiplos cromossomos). Desta forma, as células eucarióticas requerem mecanismos que garantam que uma cópia de cada cromossomo seja fielmente transmitida para cada célula nova. Entretanto, esta generalização – um único cromossomo circular nos procariotos e múltiplos cromossomos lineares nos eucariotos – nem sempre é verdadeira. Algumas bactérias têm mais de um cromossomo e importantes genes bacterianos são encontrados com frequência em outras moléculas de DNA chamadas de *Plasmídios* (ver *Capítulo 9*). Além disso, alguns genes estão localizados nas moléculas circulares de DNA encontradas em determinadas organelas de alguns eucariotos (ver *Capítulo 11*).

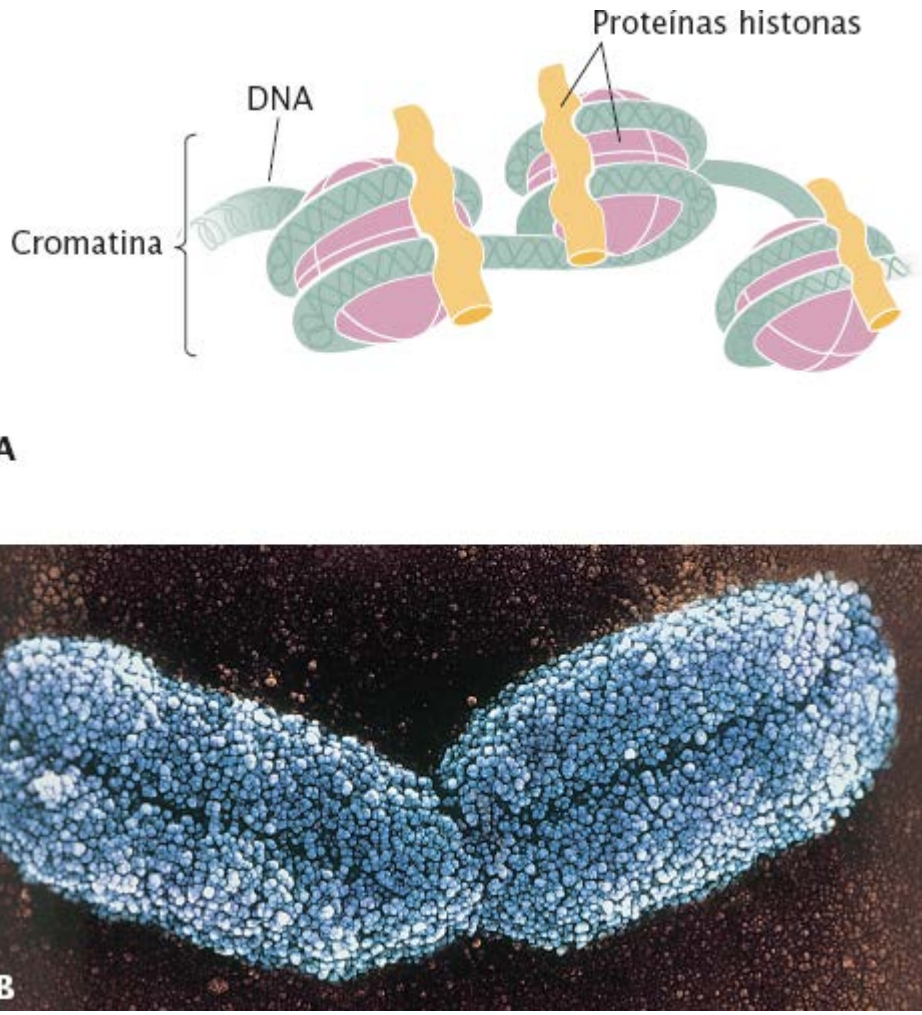


Figura 2.2 Os cromossomos dos eucariotos têm DNA e proteínas histonas. A. O DNA se dobra ao redor das histonas para formar a cromatina, o material que compõe os cromossomos. **B.** Um cromossomo eucariótico. (*Parte B*: Biophoto Associates/Science Source.)

Conceitos

Os organismos são classificados como procariotos ou eucariotos, sendo os procariotos compostos por archaea e eubactérias. Procarioto é um organismo unicelular que não tem núcleo e seu genoma é um único cromossomo. Os eucariotos são tanto unicelulares quanto multicelulares, suas células têm um núcleo, seu DNA é complexado a histonas, e seus genomas são compostos por múltiplos cromossomos.

Checagem dos conceitos 1

Liste as várias características que as eubactérias e archaea têm em comum e que as diferenciam dos eucariotos.

Os **vírus** não são procariotos nem eucariotos porque eles não têm estrutura celular. Os vírus são, na verdade, estruturas simples compostas por um revestimento proteico externo que circunda um ácido nucleico (DNA ou RNA, **Figura 2.3**). Os vírus não são as formas primitivas de vida: eles conseguem se reproduzir apenas dentro das células do hospedeiro, o que significa que eles evoluíram após as células, não antes. Além disso, os vírus não são um grupo distinto do ponto de vista evolutivo, mas estão mais relacionados com os seus hospedeiros: os genes de um vírus de planta são mais semelhantes aos de uma célula de planta do que aos genes dos vírus de animais, o que sugere que os vírus evoluíram a partir dos seus hospedeiros, em vez de outros vírus. A relação próxima entre os genes dos vírus e o hospedeiro os torna úteis para o estudo da genética dos organismos hospedeiros.

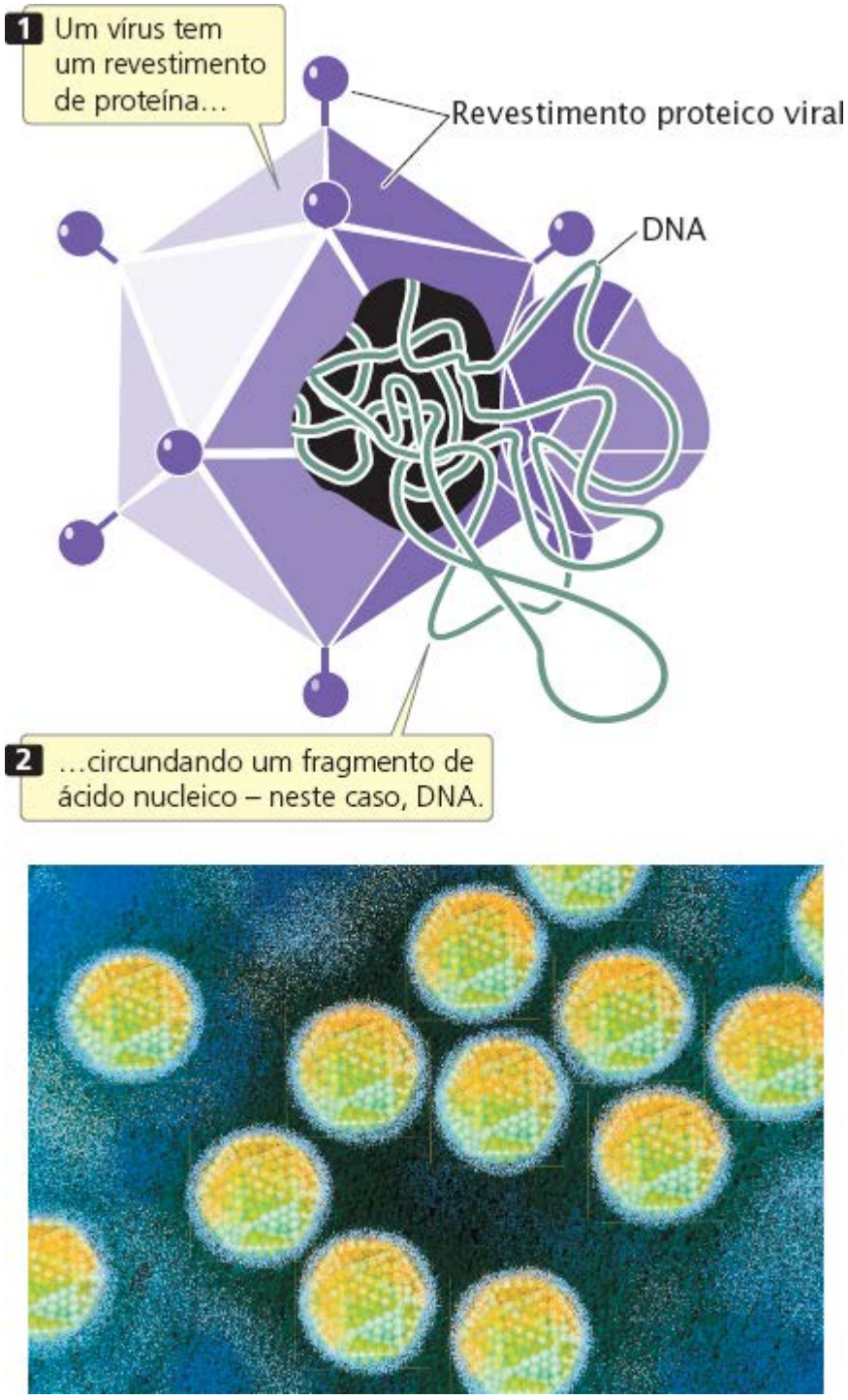


Figura 2.3 Vírus é uma estrutura de replicação simples que consiste em proteína e ácido nucleico. A micrografia apresenta adenovírus. (BSIP/Science Source.)

2.2 A reprodução celular exige a cópia do material genético, a separação das cópias e a divisão celular

Para que qualquer célula se reproduza de modo bem-sucedido, estes eventos fundamentais precisam ocorrer: (1) suas informações genéticas têm de ser copiadas, (2) as cópias das informações genéticas têm de ser separadas umas das outras e (3) a célula precisa se dividir. Toda reprodução celular inclui esses três eventos, mas os processos que levam a eles são diferentes nas células procarióticas e eucarióticas por causa das suas diferenças estruturais.

Reprodução da célula procariótica

Quando as células procarióticas se reproduzem, o cromossomo circular da bactéria se replica e a célula se divide em um processo chamado *fissão binária* (**Figura 2.4**). Em geral, a replicação começa em um lugar específico no cromossomo bacteriano, chamado de origem da replicação. Em um processo que não é totalmente compreendido, as origens de dois cromossomos recém-replicados se afastam e se deslocam no sentido dos lados opostos da célula. Pelo menos em algumas

bactérias, as proteínas se ligam próximas das origens de replicação e ancoram os novos cromossomos à membrana plasmática nos lados opostos da célula. Finalmente, uma nova parede celular é formada entre os dois cromossomos, produzindo duas células, cada uma com uma cópia idêntica dos cromossomos. Sob condições ideais, algumas células bacterianas se dividem a cada 20 min. Nesta velocidade, uma única célula bacteriana poderia produzir um bilhão de descendentes em apenas 10 h.

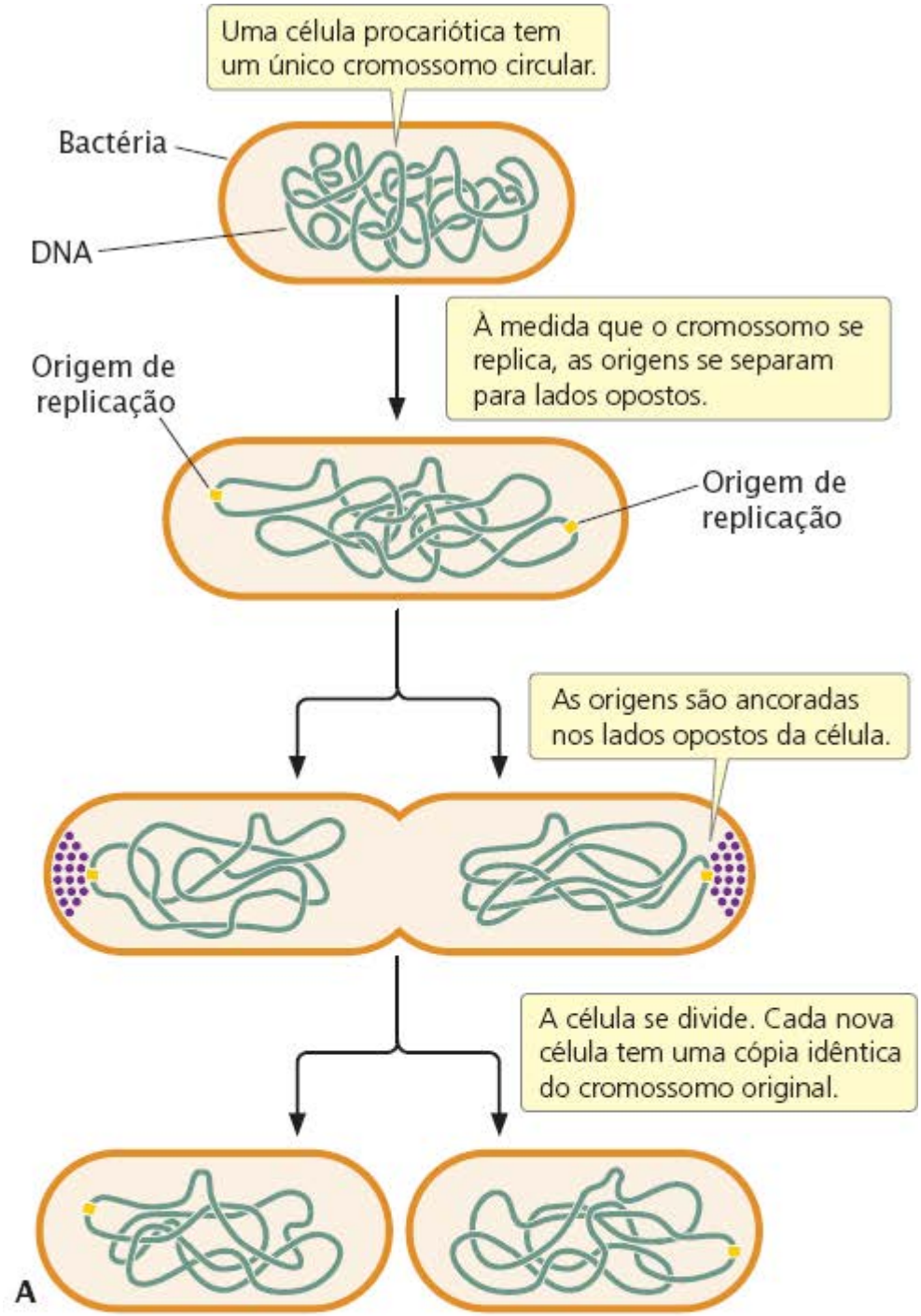
Reprodução da célula eucariótica

Como a reprodução da célula procariótica, a reprodução da célula eucariótica requer os processos de replicação de DNA, separação das cópias e divisão do citoplasma. Entretanto, a presença de múltiplas cópias da molécula de DNA requer um mecanismo mais complexo para garantir que exatamente uma cópia de cada molécula vá para cada uma das novas células.

Os cromossomos eucarióticos são separados do citoplasma pelo envelope nuclear. O núcleo tem um arcabouço interno altamente organizado, chamado de matriz nuclear – uma rede de fibras de proteínas que mantém relações espaciais precisas entre os componentes nucleares e participa da replicação do DNA, da expressão dos genes e da modificação dos produtos do gene antes de eles deixarem o núcleo. Observemos com mais detalhes a estrutura dos cromossomos eucarióticos.

Os cromossomos eucarióticos. Cada espécie eucariótica tem vários cromossomos característicos por célula: as batatas têm 48 cromossomos, as moscas-da-fruta, 8, e os humanos, 46. Parece não haver relação especial entre a complexidade de um organismo e seu número de cromossomos por célula.

Na maior parte das células eucarióticas, existem dois conjuntos de cromossomos. A presença de dois conjuntos é uma consequência da reprodução sexuada: um conjunto é herdado do genitor masculino e outro é herdado do genitor feminino. Cada cromossomo é um conjunto que tem o cromossomo correspondente no outro conjunto, constituindo juntos um par homólogo (**Figura 2.5**). As células humanas, por exemplo, têm 46 cromossomos, constituindo 23 pares homólogos.



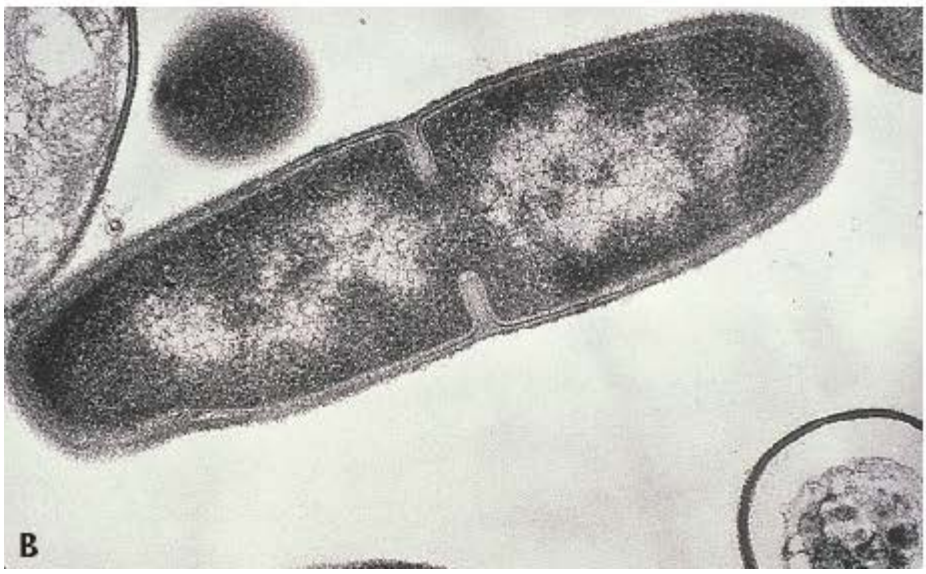


Figura 2.4 As células procarióticas se reproduzem por meio da fissão binária. A. Processo de fissão binária. **B.** Micrografia

mostrando uma célula bacteriana sofrendo fissão binária. (*Parte B*: Lee D. Simon/Photo Researchers.)

Os dois cromossomos de um **par homólogo** são geralmente semelhantes em estrutura e tamanho, e cada um carrega as informações genéticas para o mesmo conjunto de características hereditárias (os cromossomos sexuais são uma exceção e serão discutidos no Capítulo 4). Por exemplo, se um gene de determinado cromossomo codifica uma característica como a cor do cabelo, outra cópia do gene (cada cópia é chamada de alelo) na mesma posição do análogo também codifica a cor do cabelo. Entretanto, estes dois alelos não precisam ser idênticos; um pode codificar cabelo castanho, e o outro, cabelo louro. Em geral as células que carregam dois conjuntos de informações genéticas são **diploides**. A ploidia da célula indica quantos conjuntos de informações genéticas a célula tem. As células reprodutoras (como óvulos, espermatozoide e esporos) e até células não reprodutoras de alguns organismos eucariotos podem ter um único conjunto de cromossomos e seriam **haploides**. As células de alguns organismos têm mais de dois conjuntos de informações genéticas e são chamadas **poliploides** (ver *Capítulo 8*).

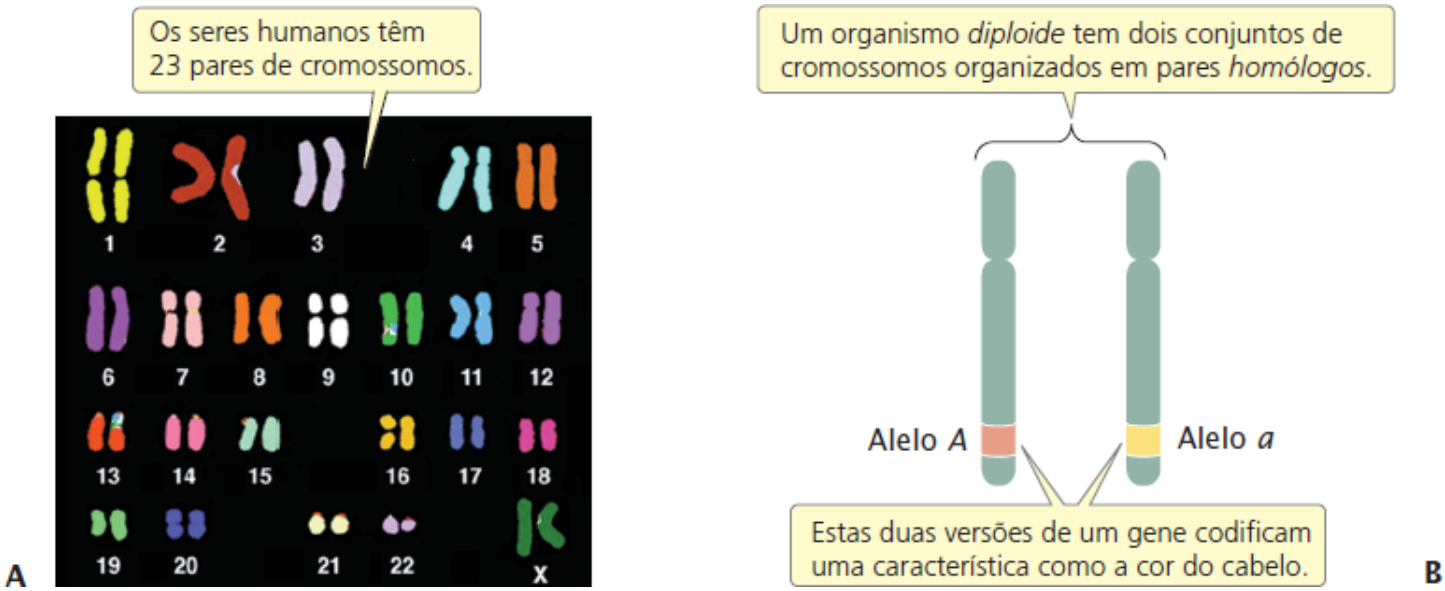


Figura 2.5 As células diploides têm dois conjuntos de cromossomos. A. Um conjunto de cromossomos a partir de uma célula humana feminina. Cada par de cromossomos é hibridizado com uma sonda de coloração singular, originando uma cor distinta. **B.** Os cromossomos estão dispostos em pares homólogos, que consistem em cromossomos com tamanho e estrutura semelhantes e carreiam informações para as mesmas características. (*Parte A*: Cortesia de Dr. Thomas Ried e Dra. Evelin Schrock.)

Conceitos

As células se reproduzem ao copiar e separar suas informações genéticas e, então, dividem-se. Como os eucariotos têm múltiplos cromossomos, existem mecanismos para garantir que cada nova célula receba uma cópia de cada cromossomo. A maior parte das células eucarióticas é diploide e seus dois conjuntos de cromossomos podem ser organizados em pares homólogos. As células haploides têm um único conjunto de cromossomos.

Checagem dos conceitos 2

As células diploides têm:

a. Dois cromossomos.
b. Dois conjuntos de cromossomos.
c. Um conjunto de cromossomos.
d. Dois pares de cromossomos homólogos.

Estrutura do cromossomo. Os cromossomos das células eucarióticas são maiores e mais complexos que os cromossomos encontrados nos procariotos, mas cada cromossomo não replicado consiste em uma única molécula de DNA. Embora sejam lineares, as moléculas de DNA nos cromossomos eucariotos estão muito dobradas e condensadas; se esticados, alguns cromossomos teriam alguns centímetros de comprimento, milhares de vezes a medida de um núcleo típico. Para acondicionar DNA desse comprimento nesse pequeno volume, cada molécula de DNA é enrolada ao redor das histonas e comprimida até formar um cromossomo em formato de bastão. Na maior parte do tempo, os cromossomos são finos e é difícil observá-los, mas antes da divisão celular, eles se condensam ainda mais e formam estruturas espessas, de fácil observação; é neste estágio que os

cromossomos são estudados.

Um cromossomo funcional tem três elementos essenciais: um centrômero, um par de telômeros e as origens de replicação. O centrômero é o ponto de fixação para os microtúbulos do fuso – os filamentos responsáveis por mover os cromossomos na divisão celular (**Figura 2.6**). O centrômero aparece como uma região contraída. Antes da divisão celular, um complexo multiproteico chamado de cinetócoro é formado no centrômero, e, posteriormente, os microtúbulos do fuso se prendem ao cinetócoro. Os cromossomos sem um centrômero não podem ser atraídos para os núcleos recém-formados; por isso, eles são perdidos, em muitos casos com consequências catastróficas para a célula. Com base na localização do centrômero, os cromossomos são classificados em quatro tipos: metacêntricos, submetacêntricos, acrocêntricos e telocêntricos (**Figura 2.7**).

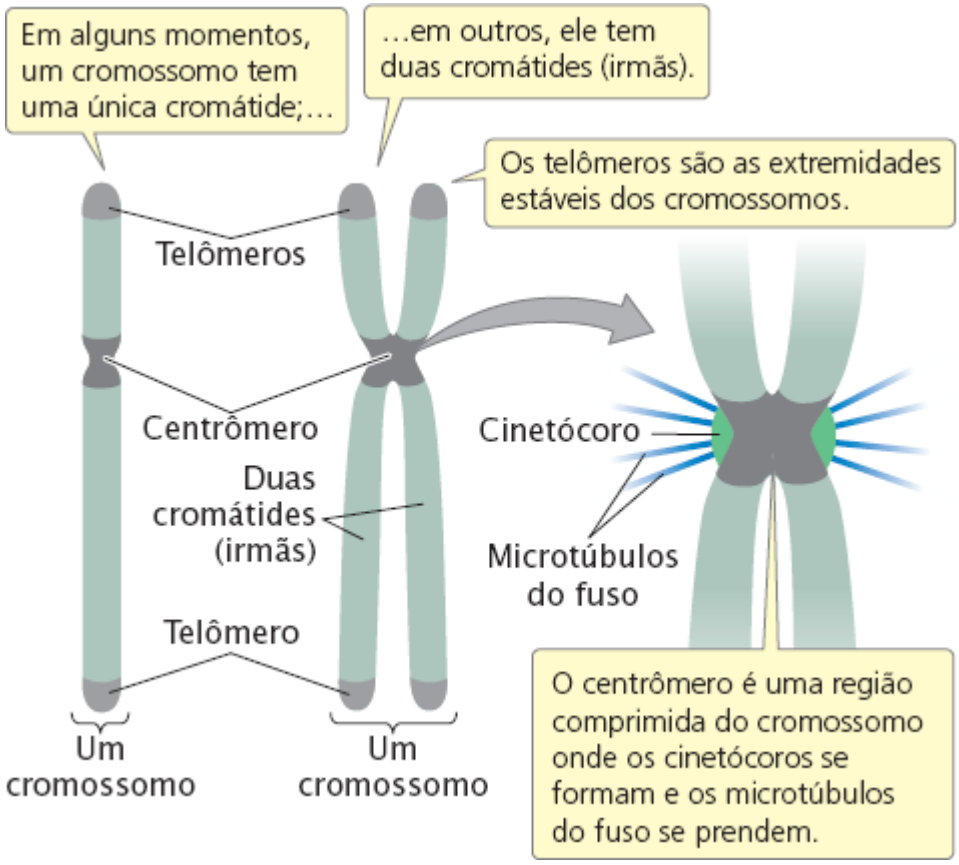


Figura 2.6 Cada cromossomo eucariótico tem um centrômero e telômeros.

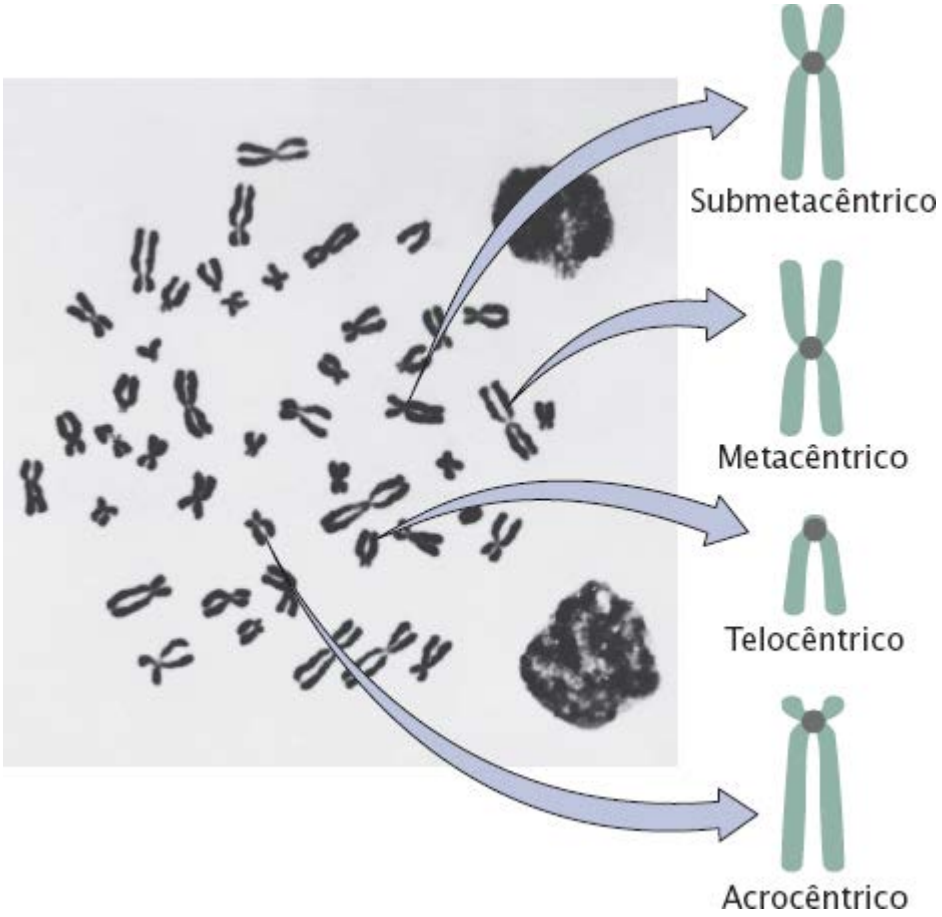


Figura 2.7 Existem quatro tipos principais de cromossomos eucarióticos dependendo da posição do centrômero. (Micrografia de Don W. Fawcett/Science Source.)

Os **telômeros** são as extremidades naturais, as pontas, de um cromossomo linear completo (ver Figura 2.6). Como as pontas de plástico que protegem as pontas de um cadarço, os telômeros protegem e estabilizam as extremidades dos cromossomos. Se um cromossomo se romper e produzir novas extremidades, ele é degradado nas extremidades recém-quebradas. Os telômeros estabilizam o cromossomo. A pesquisa mostra que os telômeros também participam na limitação da divisão celular e são importantes no envelhecimento e no câncer (*Capítulo 12*).

As **origens de replicação** são locais onde começa a síntese de DNA, e não é fácil observá-las no microscópio. Sua estrutura e função serão discutidas com mais detalhes no Capítulo 12. Na preparação para divisão celular, cada cromossomo se replica, fazendo uma cópia de si mesmo, como já foi mencionado. Estas duas cópias inicialmente idênticas, chamadas de **cromátides-irmãs**, são mantidas juntas no centrômero (**Figura 2.6**). Cada cromátide-irmã tem uma única molécula de DNA.

Conceitos

As cromátides-irmãs são cópias de um cromossomo unido pelo centrômero. Os cromossomos funcionais têm centrômeros, telômeros e origens de replicação. O cinetócoro é o ponto de fixação para os microtúbulos do fuso, os telômeros são extremidades de estabilização de um cromossomo, as origens de replicação são locais onde começa a síntese do DNA.

Checagem dos conceitos 3

O que aconteceria se um cromossomo não tivesse um cinetócoro?

O ciclo celular e a mitose

O **ciclo celular** é a história de vida de uma célula, os estágios pelos quais ela passa de uma divisão para a próxima (**Figura 2.8**). Este processo é crítico para a genética, porque, por meio do ciclo celular, as instruções genéticas para todas as características são transmitidas da célula-mãe para as células-filhas. Um novo ciclo inicia após uma célula ter se dividido e produzido duas novas células. Cada nova célula metaboliza, cresce e se desenvolve. No final de seu ciclo, a célula se divide para produzir duas células, que podem então sofrer ciclos celulares adicionais. A progressão durante o ciclo celular é regulada em pontos-chave de transição, chamados de **pontos de verificação**, que permitem ou proíbem a progressão da célula para o próximo estágio. Os

pontos de verificação garantem que todos os componentes celulares estejam presentes e funcionais e são necessários para evitar que células com cromossomos danificados ou ausentes se proliferem. Defeitos nos pontos de verificação podem resultar em crescimento celular desregulado, como observado em alguns cânceres. A base molecular desses pontos será discutida no Capítulo *23*.

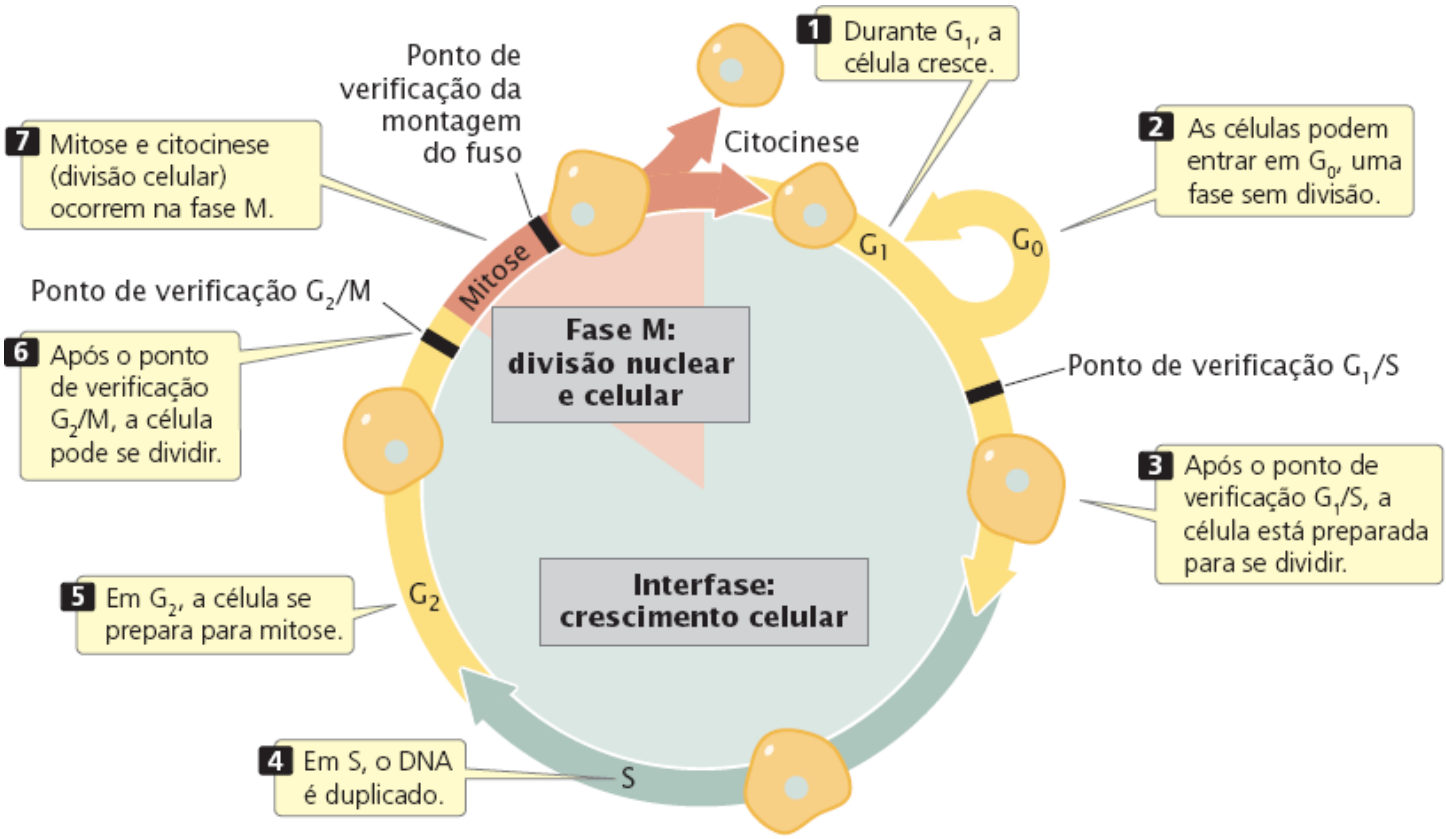


Figura 2.8 O ciclo celular é composto pela interfase e fase M.

O ciclo celular tem duas fases principais. A primeira é a **interfase**, o período entre as divisões celulares, na qual a célula cresce, desenvolve-se e funciona. Na interfase, também ocorrem eventos críticos necessários para a divisão celular. A segunda principal fase é a **fase M (mitótica)**, o período de divisão celular ativo. A fase M inclui a **mitose**, o processo de divisão nuclear e a **citocinese** ou divisão do citoplasma. Vamos observar os detalhes da interfase e da fase M.

Interfase. A interfase é o período estendido de crescimento e desenvolvimento entre as divisões celulares. Embora possa ser observada pouca atividade com um microscópio óptico, a célula está bem ocupada: o DNA está sendo sintetizado, o RNA e as proteínas estão sendo produzidos e milhares de reações bioquímicas necessárias para as funções celulares estão ocorrendo. Além do crescimento e desenvolvimento, a interfase inclui vários pontos de verificação.

Por convenção, a interfase é dividida em três subfases: G_1, S e G_2 (ver Figura 2.8). A interfase começa com G_1 (para *gap* – intervalo – 1). Em G_1 a célula cresce e as proteínas necessárias para a divisão celular são produzidas; em geral esta fase dura algumas horas. Próximo do término de G_1, um ponto crítico, chamado de ponto de verificação G_1/S, mantém a célula em G_1 até ela ter todas as enzimas necessárias para a replicação do DNA. Após este ponto de verificação passar, a célula está preparada para se dividir.

Antes de alcançar o ponto de verificação de G_1/S, as células podem sair do ciclo celular ativo em resposta a sinais regulatórios e passar para uma fase de não divisão chamada de G_0, que é um estado estável no qual as células em geral mantêm um tamanho constante. Elas podem permanecer na G_0 por um bom tempo, até indefinidamente, ou podem entrar novamente na G_1 e no ciclo celular ativo. Muitas células nunca entram em G_0; ao contrário, elas têm um ciclo contínuo.

Após G_1, a célula entra na fase S (para síntese de DNA) na qual cada cromossomo é duplicado. Embora a célula esteja empenhada em se dividir após ter passado pelo ponto de verificação G1/S, a síntese de DNA pode ocorrer antes que a célula possa ir para a fase de mitose. Se a síntese de DNA é bloqueada (por fármacos ou uma mutação), a célula não será capaz de sofrer mitose. Antes da fase S, cada cromossomo não é replicado; após a fase S, cada cromossomo é composto por duas cromátides **(**ver Figura 2.6).

Após a fase S, a célula entra na fase G_2 (*gap* – intervalo – 2). Nesta fase, ocorrem vários eventos bioquímicos adicionais que são necessários para a divisão celular. O importante ponto de verificação G_2/M é alcançado próximo do final de G_2. Este ponto

é ultrapassado apenas se o DNA da célula estiver completamente replicado e íntegro. O DNA não replicado ou danificado pode inibir a ativação de algumas proteínas necessárias para que ocorra mitose. Após passar pelo ponto de verificação G_2/M, a célula está pronta para se dividir e entra na fase M. Embora a duração da interfase varie dependendo do tipo de célula, uma típica célula de mamífero em divisão leva cerca de 10 h em G_1, 9 h em S e 4 h em G_2 (ver Figura 2.8).

Durante a interfase, os cromossomos estão em um estado relaxado, mas não desenrolado e os cromossomos individuais não podem ser observados com um microscópio. Esta condição muda drasticamente quando a interfase chega ao final e a célula entra na fase M (mitótica).

Fase M. A fase M é a parte do ciclo celular na qual as cópias dos cromossomos da célula (as cromátides-irmãs) se separam e a célula sofre divisão. A separação das cromátides-irmãs na fase M é um processo crítico que resulta em um conjunto completo de informações genéticas para cada uma das células resultantes. Os biólogos geralmente dividem a fase M em seis estágios: os cinco estágios de mitose (prófase, pró-metáfase, metáfase, anáfase e telófase), ilustrados na **Figura 2.9**, e citocinese. É importante ter em mente que a fase M é um processo contínuo e que sua separação nesses seis estágios é um pouco arbitrária.

Prófase. À medida que a célula entra na prófase, os cromossomos se tornam visíveis no microscópio óptico. Como o cromossomo foi duplicado na fase S anterior, cada cromossomo tem duas cromátides presas no centrômero. O fuso mitótico, um arranjo organizado de microtúbulos que movem os cromossomos na mitose, forma-se. Nas células animais, o fuso cresce a partir dos centrossomos que migram para os lados opostos da célula. Dentro de cada cromossomo existe uma organela especial, o centríolo, que também é composto por microtúbulos. Algumas células de plantas não têm centrossomos nem centríolos, mas têm fusos mitóticos.

Pró-metáfase. A desintegração da membrana celular marca o início da pró-metáfase. Os microtúbulos do fuso, que estavam fora do núcleo, entram na região nuclear e são compostos por subunidades de uma proteína chamada tubulina (**Figura 2.10**). Durante a pró-metáfase, as moléculas de tubulina são adicionadas e removidas dos microtúbulos, fazendo com que eles sofram ciclos repetidos de crescimento e contração. O microtúbulo é estabilizado quando sua extremidade encontra um cinetócoro. Finalmente, cada cromossomo é preso aos microtúbulos a partir dos polos opostos do fuso: para cada cromossomo, um microtúbulo de um dos centrossomos se ancora ao cinetócoro de uma das cromátides-irmãs; um microtúbulo do centrossomo oposto então se prende a outra cromátide-irmã, ancorando o cromossomo para ambos os centrossomos. Essa organização é conhecida como biorientação do cromossomo.

Metáfase. Durante a metáfase, os cromossomos são organizados em um único plano, a placa equatorial, entre os dois centrossomos. Estes, agora em extremidades opostas da célula com microtúbulos irradiando para o extremo e se encontrando no centro da célula, centralizam-se nos polos do fuso. Um ponto de verificação da montagem dos polos garante que cada cromossomo esteja alinhado na placa equatorial e preso a fibras do fuso dos polos opostos.

A passagem de uma célula pelo ponto de verificação da montagem do fuso depende da tensão gerada no cinetócoro à medida que as duas cromátides unidas são puxadas em sentidos opostos pelas fibras do fuso. Essa tensão é necessária para que a célula passe pelo ponto de verificação da montagem do fuso. Se um microtúbulo se fixa a uma cromátide, mas não a outra, não é gerada tensão e a célula não é capaz de progredir para o próximo estágio do ciclo celular. O ponto de verificação da montagem do fuso é capaz de detectar até um único par de cromossomos que não esteja adequadamente fixado aos microtúbulos. A importância desse ponto é ilustrada pelas células que têm problemas no seu ponto de verificação da montagem do fuso, quando, muitas vezes, terminam com um número anormal de cromossomos.

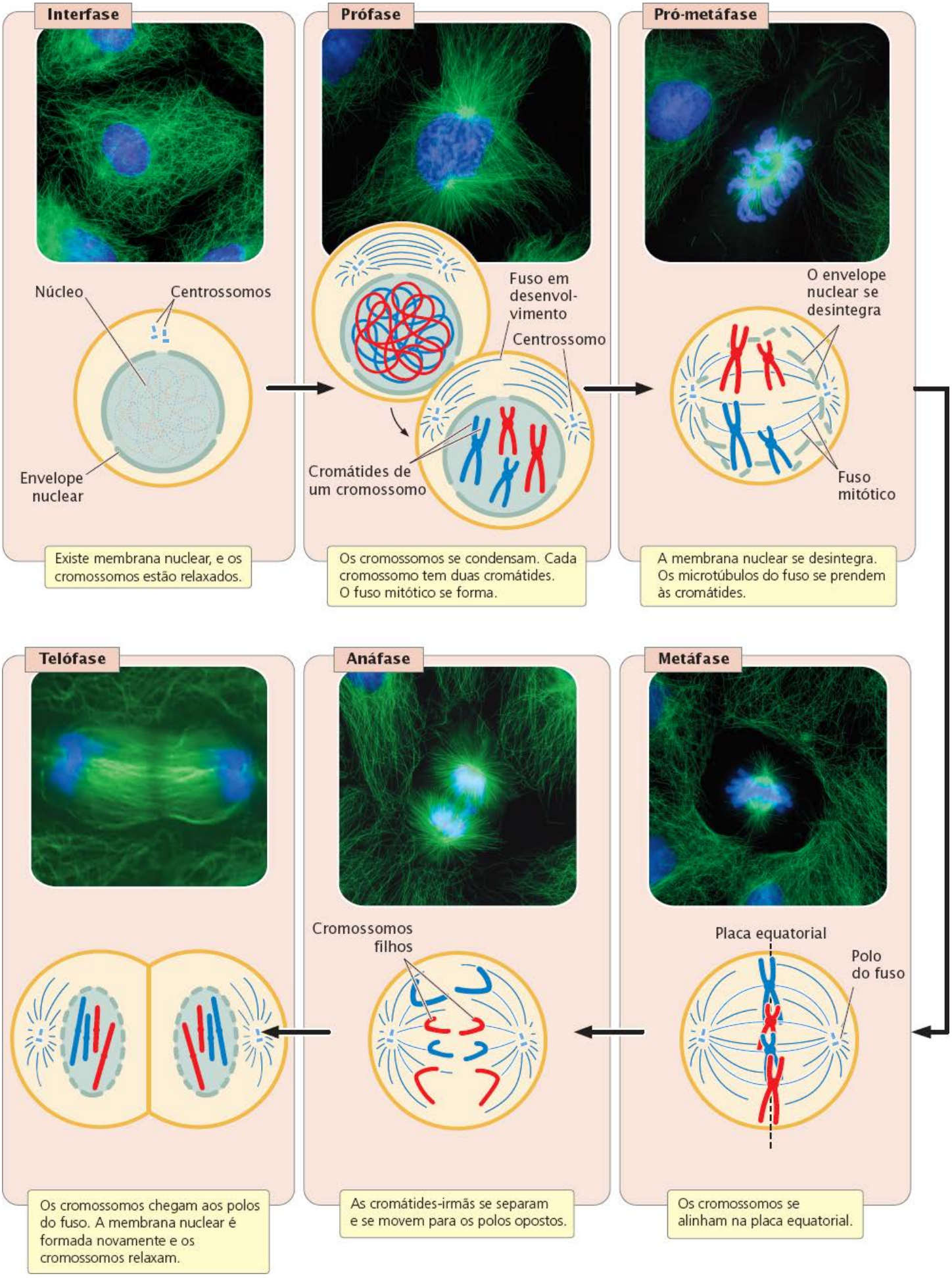


Figura 2.9 O ciclo celular é dividido em estágios. (Fotografias de Conly L. Rieder/Biological Photo Service.)

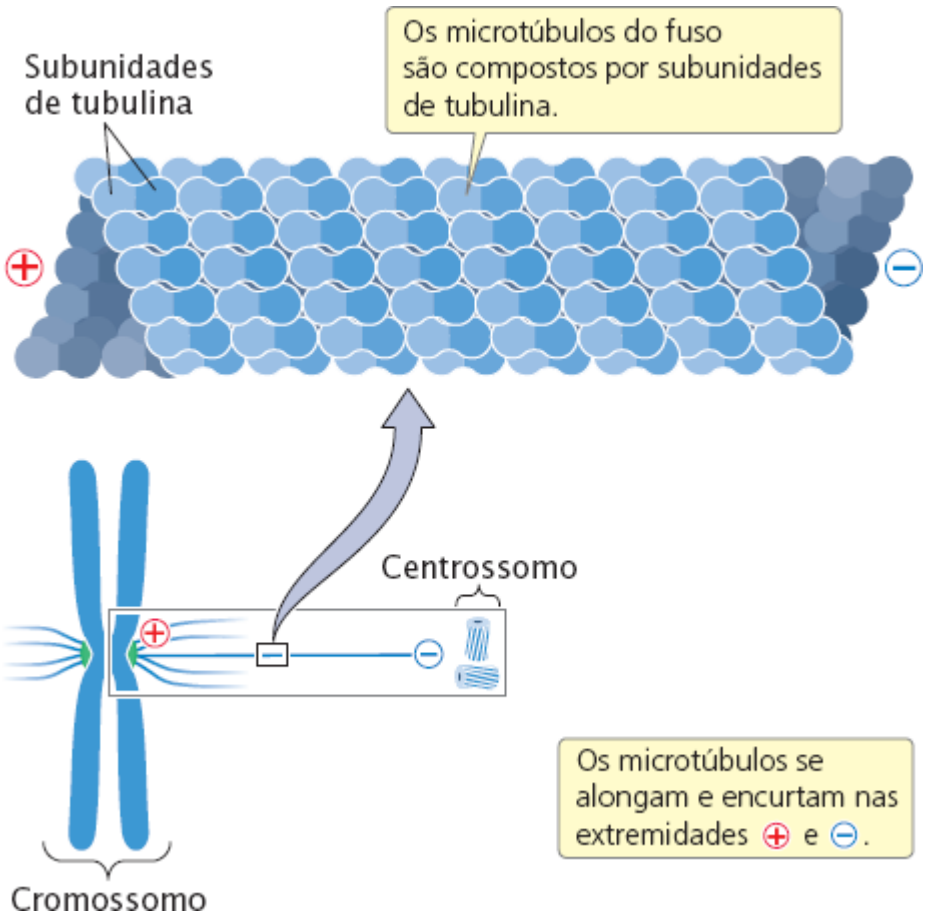


Figura 2.10 Os microtúbulos são compostos por subunidades de tubulina. Cada microtúbulo tem uma extremidade positiva (+) no cinetócoro e sua extremidade negativa (–) no final do cromossomo.

Anáfase. Após ter passado pelo ponto de verificação da montagem do fuso, a conexão entre as cromátides-irmãs é rompida e elas se separam. Isso marca o início da anáfase, durante a qual os cromossomos se deslocam no sentido dos polos opostos do fuso. O movimento do fuso ocorre graças à desmontagem das moléculas de tubulina na extremidade do cinetócoro (chamada de extremidade +) e a extremidade do fuso (chamada de extremidade −) da fibra do fuso (ver Figura 2.10). Proteínas especiais chamadas de motores moleculares desmontam as moléculas de tubulina do fuso e geram forças que puxam o cromossomo para o polo do fuso.

Telófase. Após as cromátides terem se separado, cada uma é considerada um cromossomo. A telófase é marcada pela chegada dos cromossomos nos polos do fuso. A membrana nuclear é formada novamente ao redor de cada conjunto de cromossomos, produzindo dois núcleos separados dentro da célula. Os cromossomos relaxam e se alongam, mais uma vez desaparecendo do campo visual. Em muitas células, a divisão do citoplasma (citocinese) ocorre junto com a telófase.

As principais características do ciclo celular estão resumidas no **Quadro 2.1**. Resolva o Problema 23

Consequências genéticas do ciclo celular

Quais são os resultados importantes do ponto de vista genético do ciclo celular? A partir de uma única célula, o ciclo celular produz duas células que têm as mesmas instruções genéticas. As células-filhas resultantes são geneticamente idênticas uma à outra e à sua célula genitora porque a síntese do DNA na fase S cria uma cópia exata de cada molécula de DNA, dando origem a duas cromátides-irmãs idênticas. A mitose garante que uma das duas cromátides-irmãs de cada cromossomo replicado passe para cada nova célula.

Quadro 2.1 Características do ciclo celular.

Estágio	Principais características
Fase G_0	Período estável, de não divisão e duração variável
Interfase	
Fase G_1	Crescimento e desenvolvimento da célula; ponto de verificação G_1/S

Fase S	Síntese de DNA
Fase G_2	Preparação para divisão; ponto de verificação G_2/M
Fase M	
Prófase	Os cromossomos se condensam e os fusos mitóticos se formam
Pró-metáfase	O envelope nuclear se desintegra e os microtúbulos do fuso se ancoram nos cinetócoros
Metáfase	Os cromossomos se alinham na placa equatorial; ponto de verificação da montagem do fuso
Anáfase	As cromátides-irmãs se separam e se tornam cromossomos individuais que migram na direção dos polos do fuso
Telófase	Os cromossomos chegam aos polos do fuso, o envelope nuclear se forma novamente e os cromossomos condensados relaxam
Citocinese	O citoplasma se divide; a parede celular se forma nas células vegetais

Outro resultado importante do ponto de vista genético do ciclo celular é que cada uma das células produzidas tem um complemento total dos cromossomos: não há redução ou aumento no número de cromossomos. Cada célula também tem aproximadamente metade do citoplasma e organelas da célula genitora original, mas nenhum mecanismo preciso e homólogo para mitose garante que as células sejam divididas exatamente. Consequentemente, nem todas as células resultantes do ciclo celular têm material citoplasmático idêntico.

Conceitos

As fases ativas do ciclo celular são a interfase e a fase M. A interfase consiste em G_1, S e G_2. Em G_1, a célula cresce e se prepara para a divisão celular; na fase S, ocorre a síntese de DNA; e em G_2, ocorrem outros eventos bioquímicos necessários para a divisão celular. Algumas células entram em uma fase quiescente chamada de G_0. A fase M inclui a mitose e a citocinese e é dividida em prófase, pró-metáfase, metáfase, anáfase e telófase. O ciclo celular produz duas células geneticamente idênticas, e cada uma tem um conjunto completo de cromossomos.

Checagem dos conceitos 4

Qual é a ordem correta dos estágios do ciclo celular?

a. G_1, S, prófase, metáfase, anáfase.
b. S, G_1, prófase, metáfase, anáfase.
c. Prófase, S, G_1, metáfase, anáfase.
d. S, G_1, anáfase, prófase, metáfase.

Conceitos conectantes

Contagem de cromossomos e moléculas de DNA

As relações entre cromossomos, cromátides e moléculas de DNA causam confusão com muita frequência. Em alguns momentos, os cromossomos não são replicados, em outros, cada um tem duas cromátides (ver Figura 2.6). Às vezes, os cromossomos são uma única molécula de DNA, em outros casos, eles têm duas moléculas de DNA. Como podemos rastrear o número dessas estruturas no ciclo celular?

Existem duas regras simples para contar os cromossomos e as moléculas de DNA: (1) para determinar o número de cromossomos, contar o número de centrômeros funcionais; (2) para determinar o número de moléculas de DNA, determinar primeiro se existem cromátides-irmãs. Caso existam, o cromossomo se replicou e o número de moléculas de DNA é o dobro do número de cromossomos. Se não houver cromátides-irmãs, o cromossomo não se replicou e o número de moléculas de DNA é o mesmo dos cromossomos.

Vamos examinar uma célula hipotética à medida que ela passa pelo ciclo celular (**Figura 2.11**). No início de G_1, esta célula diploide tem dois conjuntos completos de cromossomos, para um total de quatro cromossomos. Cada cromossomo não está replicado e consiste em uma única molécula de DNA, existindo quatro moléculas de DNA na célula durante G_1. Cada molécula é copiada na fase S. As duas moléculas de DNA resultantes se combinam com as histonas e outras proteínas para formar cromátides-irmãs. Embora a quantidade de DNA dobre na fase S, o número de cromossomos permanece o mesmo porque as cromátides-irmãs estão presas e compartilham um único centrômero funcional. Ao término da fase S, esta célula ainda tem quatro cromossomos, cada um com duas cromátides-irmãs; então, 4 × 2 = 8 moléculas de DNA estão presentes.

Do início ao fim da prófase, pró-metáfase e metáfase, a célula tem quatro cromossomos e oito moléculas de DNA. Entretanto, as cromátides-irmãs se separam na anáfase. Cada uma agora tem seu próprio centrômero funcional, e cada uma é considerada um cromossomo separado. Até a citocinese, a célula tem oito cromossomos não replicados, ainda existindo oito moléculas de DNA presentes. Após a citocinese, os oito cromossomos (e as oito moléculas de DNA) estão igualmente distribuídos entre as duas células; então cada nova célula tem quatro cromossomos e quatro moléculas de DNA, o número presente no início do ciclo celular.

Em resumo, o número de cromossomos aumenta apenas na anáfase, quando as duas cromátides de um cromossomo se separam e se tornam cromossomos distintos. O número de cromossomos reduz apenas na citocinese. O número de moléculas de DNA aumenta apenas na fase S e reduz apenas na citocinese.

Resolva o Problema 26

2.3 A reprodução sexuada produz variação genética pela meiose

Se toda reprodução fosse feita pela mitose, a vida seria bem entediante, porque a mitose produz apenas células geneticamente idênticas. Com a mitose apenas, você, seus filhos, seus pais, irmãos e irmãs, sobrinhos, e muitas pessoas que você nem conhece seriam clones – cópias um do outro. Apenas a mutação eventual introduziria variabilidade genética. Todos os organismos se reproduziram desta forma pelos primeiros 2 bilhões de anos de existência da Terra (e é a forma pela qual alguns organismos ainda se reproduzem hoje). Então, cerca de 1,5 a 2 bilhões de anos atrás, algo incrível evoluiu: as células que produzem descendentes geneticamente variáveis por meio da reprodução sexuada.

A evolução da reprodução sexuada está entre os eventos mais importantes na história da vida. Como será discutido nos Capítulos 24 e 25, o ritmo da evolução depende da variação genética. Ao misturar as informações genéticas dos pais, a reprodução sexuada aumenta muito a variação genética e possibilita a aceleração da evolução. A maior parte da grande diversidade de vida na Terra é resultado direto da reprodução sexuada.

A reprodução sexuada tem dois processos. O primeiro é a **meiose**, que gera os gametas, nos quais o número de cromossomos é reduzido à metade. O segundo processo é a **fertilização**, na qual dois gametas haploides se fundem e restauram o número de cromossomos para sua forma diploide original.

Meiose

Às vezes os termos *mitose* e *meiose* são confundidos. Eles são semelhantes e ambos se referem à divisão de cromossomo e à citocinese. Mas não se engane. Os desfechos da mitose e meiose são radicalmente diferentes e vários eventos únicos que têm consequências genéticas importantes ocorrem apenas na meiose.

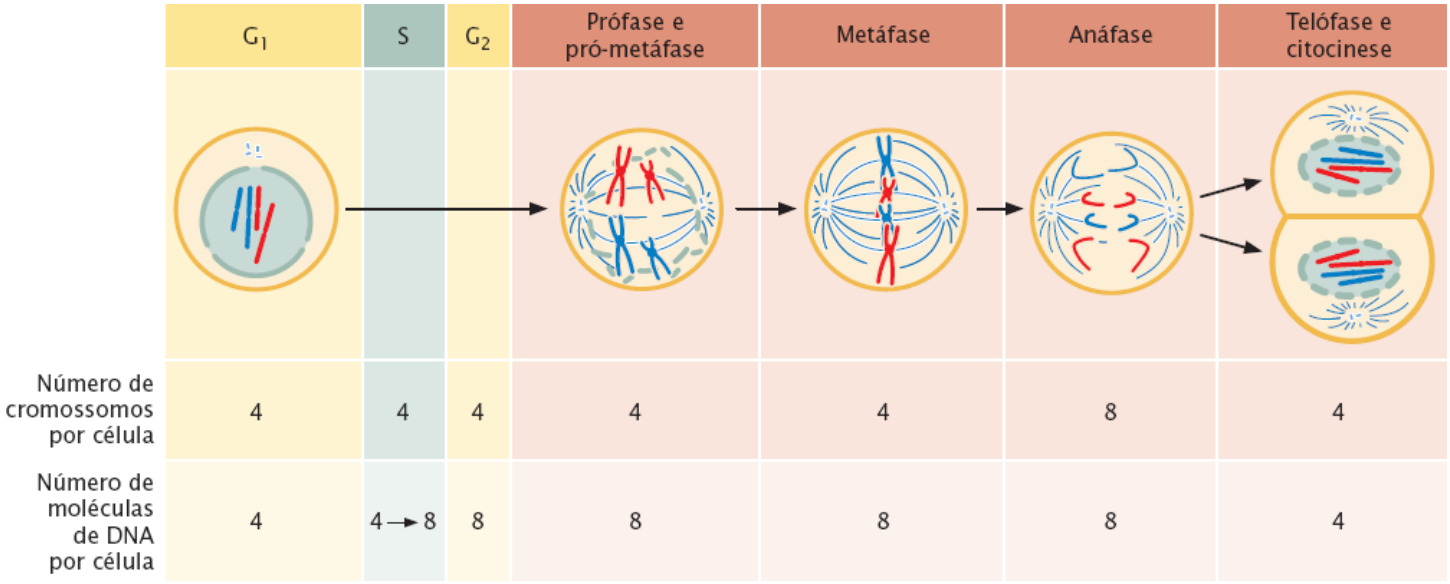


Figura 2.11 O número de cromossomos e o número de moléculas de DNA mudam durante o curso do ciclo celular. O número de cromossomos por célula é igual ao número de centrômeros funcionais. O número de moléculas de DNA por célula é igual ao número de cromossomos quando estes não estão replicados (sem cromátides-irmãs presentes) e o dobro do número de cromossomos quando *existem* cromátides-irmãs.

Qual a diferença entre meiose e mitose? A mitose consiste em uma única divisão nuclear e geralmente é acompanhada por uma única divisão celular. Na meiose, por outro lado, ocorrem duas divisões. Após a mitose, o número de cromossomos nas células recém-formadas é o mesmo que na célula original, enquanto na meiose o número de cromossomos nas células recém-formadas é a metade. Finalmente, a mitose produz células geneticamente idênticas, enquanto as células são diferentes do ponto de vista genético na meiose. Vejamos como essas diferenças surgem.

Como a mitose, a meiose é precedida por um estágio de interfase que inclui as fases G_1, S e G_2. A meiose tem dois processos distintos: meiose I e meiose II, cada qual incluindo uma divisão celular. A primeira, que ocorre no final da meiose I, é chamada de divisão reducional porque o número de cromossomos por célula é reduzido à metade (**Figura 2.12**). A segunda divisão, que ocorre no final da meiose II, é chamada de divisão equacional. Os eventos que ocorrem na meiose II são semelhantes aos que ocorrem na mitose. Entretanto, a meiose II difere da mitose porque o número de cromossomo já foi dividido à metade na meiose I e a célula não começa com o mesmo número de cromossomos como a célula na mitose (ver Figura 2.12).

Meiose I. Durante a interfase, os cromossomos estão relaxados e visíveis como cromatina difusa. A **prófase I** é um estágio prolongado, dividido em cinco subestágios (**Figura 2.13**). No leptóteno, os cromossomos se contraem e se tornam visíveis. No zigóteno, os cromossomos continuam a se condensar, os cromossomos homólogos pareiam, e começa a **sinapse**, um pareamento muito próximo. Cada par homólogo de cromossomos que sofreram sinapse consiste em quatro cromátides chamadas de **bivalentes** ou **tétrades**. No paquíteno, os cromossomos se tornam menores e mais espessos, e o complexo sinaptonêmico de três partes se desenvolve entre os cromossomos homólogos. A função do complexo sinaptonêmico não está clara, mas os cromossomos de muitas células que não têm este complexo não se separam corretamente.

O ***crossing over*** **(permutação)** ocorre na prófase I, na qual os cromossomos homólogos trocam informações genéticas. Ele é responsável pela variação genética (ver *Fontes de variação genética na meiose*, posteriormente) e é essencial para o alinhamento e separação adequados dos cromossomos homólogos. Os centrômeros dos cromossomos pareados se separam no diplóteno; os dois homólogos permanecem presos no mesmo quiasma, que é o resultado do *crossing over*. Próximo do final da prófase I, a membrana nuclear se rompe e o fuso se forma, preparando o estágio para a metáfase I. Os estágios da meiose estão destacados na **Figura 2.14**.

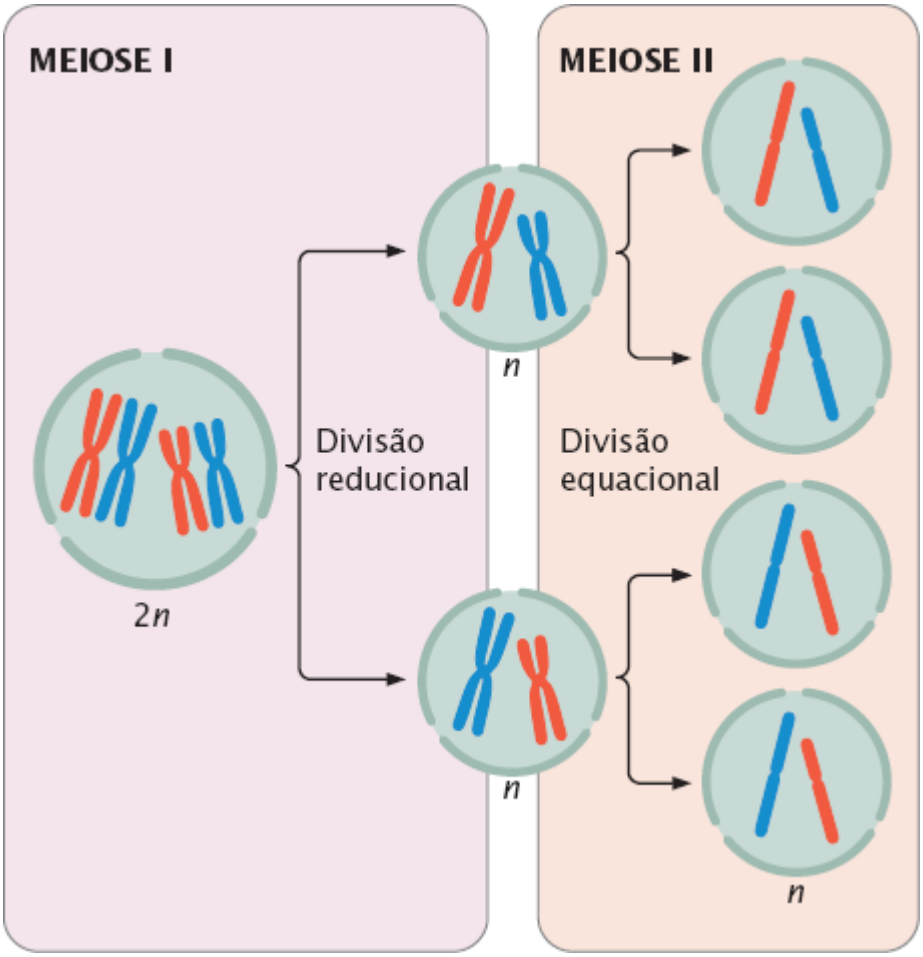


Figura 2.12 A meiose inclui duas divisões celulares. Nesta ilustração, a célula original é 2*n* = 4. Após duas divisões meióticas, cada célula resultante é 1*n* = 2.

A **metáfase I** é iniciada quando os pares homólogos dos cromossomos se alinham ao longo da placa de metáfase (ver Figura 2.14). Um microtúbulo de um polo se fixa a um cromossomo de um par homólogo e um microtúbulo do outro polo se fixa ao outro membro do par. A **anáfase I** é marcada pela separação dos cromossomos homólogos. Os dois cromossomos de um par homólogo são puxados para os polos opostos. Embora os cromossomos homólogos se separem, as cromátides-irmãs permanecem presas e se deslocam juntas. Na **telófase I**, os cromossomos chegam aos polos do fuso e o citoplasma se divide.

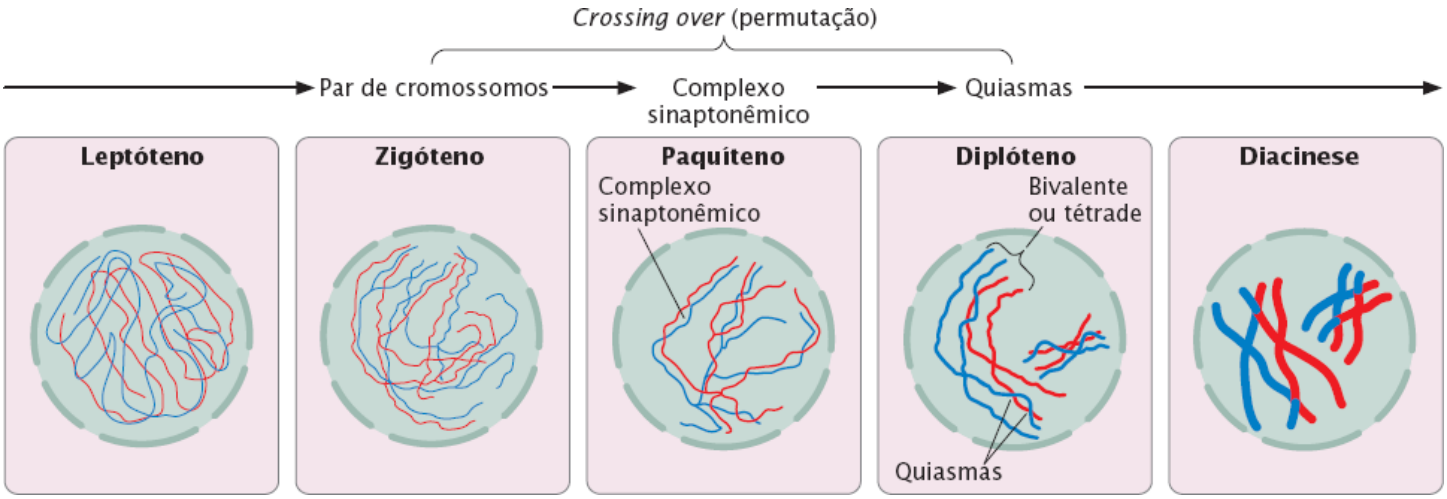


Figura 2.13 O *crossing over* (permutação) ocorre na prófase I.

Meiose II. O período entre a meiose I e meiose II é a **intercinese**, na qual a membrana nuclear se forma novamente ao redor dos cromossomos agrupados em cada polo, o fuso se rompe e os cromossomos relaxam. Essas células, então, passam para a **prófase II**, na qual os eventos na intercinese são invertidos, os cromossomos se recondensam, o fuso se forma novamente, e o envelope nuclear novamente se rompe. Na intercinese, em alguns tipos de células, os cromossomos permanecem condensados e o fuso não se desfaz. Essas células passam diretamente da citocinese para a **metáfase II**, que é semelhante à metáfase da mitose:

os cromossomos individuais se alinham com a placa de metáfase, com as cromátides-irmãs nos polos opostos.

Na **anáfase II**, os cinetócoros das cromátides-irmãs se separam e as cromátides são puxadas para os polos opostos. Cada cromátide agora é um cromossomo distinto. Na **telófase II**, os cromossomos chegam aos polos do fuso, um envelope nuclear se forma ao redor dos cromossomos e o citoplasma se divide. Os cromossomos se dividem e não são mais visíveis.

Os principais eventos da meiose estão resumidos no **Quadro 2.2**.

Quadro 2.2 Principais eventos em cada estágio da meiose.

Estágio	Principais características
Meiose I	
Prófase I	Os cromossomos se condensam, os cromossomos homólogos pareiam, ocorre o *crossing over* (permutação), o envelope nuclear se rompe e o fuso mitótico é formado
Metáfase I	Os pares de cromossomos homólogos se alinham na placa equatorial
Anáfase I	Os dois cromossomos (cada um com duas cromátides) de cada par homólogo se separam e se movem para os polos opostos
Telófase I	Os cromossomos chegam aos polos do fuso
Citocinese	O citoplasma se divide para produzir duas células, cada uma tendo metade do número original de cromossomos
Intercinese	Em alguns tipos de células, o fuso se rompe, os cromossomos relaxam e o envelope nuclear se forma novamente, mas não ocorre síntese de DNA
Meiose II	
Prófase II*	Os cromossomos se condensam, o fuso se forma e o envelope nuclear se desintegra
Metáfase II	Os cromossomos individuais se alinham na placa equatorial
Anáfase II	As cromátides-irmãs se separam e se movem como cromossomos individuais para os polos do fuso
Telófase II	Os cromossomos chegam aos polos do fuso, o fuso se rompe e um novo envelope nuclear é formado
Citocinese	O citoplasma se divide

*Apenas nas células em que o fuso se rompeu, os cromossomos relaxaram e o envelope nuclear foi formado novamente na telófase I. Outros tipos de células seguem diretamente para a metáfase II após a citocinese.

Conceitos

A meiose consiste em dois processos distintos: meiose I e meiose II. A meiose I inclui a divisão reducional, na qual os cromossomos homólogos se separam e o número de cromossomos é dividido à metade; As cromátides se separam na meiose II (a divisão equacional).

Checagem dos conceitos 5

Qual dos eventos abaixo ocorre na metáfase I?

a. *Crossing over.*
b. Os cromossomos se contraem.
c. Os pares de cromossomos homólogos se alinham na placa equatorial.
d. Os cromossomos individuais se alinham na placa equatorial.

Fontes de variação genética na meiose

Quais são as consequências gerais da meiose? Primeiro, a meiose inclui duas divisões; então, cada célula original produz quatro células (existem exceções a esta generalização, como em muitas fêmeas; ver Figura 2.19 B). Segundo, o número de cromossomos é dividido à metade, então as células produzidas pela meiose são haploides. Terceiro, as células produzidas por meiose são geneticamente diferentes uma da outra e da célula original. As diferenças genéticas entre as células resultantes destes dois processos são únicas para a meiose: o *crossing over* e a separação aleatória dos cromossomos homólogos.

Crossing over. O *crossing over*, que ocorre na prófase I, refere-se à troca de material genético entre cromátides não irmãs (cromátides oriundas de diferentes cromossomos homólogos). Evidências a partir de levedura sugerem que o *crossing over* é iniciado no zigóteno, antes de desenvolver o complexo sinaptonêmico e não está completo até perto do final da prófase I (Figura 2.13). Em outros organismos, o *crossing over* é iniciado após a formação do complexo sinaptonêmico e em outros não há complexo sinaptonêmico.

As cromátides-irmãs não são mais idênticas após o *crossing over*. Ele é a base para **recombinação** intracromossômica, criando novas combinações de alelos em uma cromátide. Para observar como o *crossing over* produz a variação genética, considere dois pares de alelos, que vamos abreviar como *Aa* e *Bb*. Imagine que um cromossomo tem os alelos *A* e *B* e seu homólogo tem os alelos *a* e *b* (**Figura 2.15 A**). Quando o DNA é replicado na fase S, cada cromossomo é duplicado, e então as cromátides-irmãs são idênticas (**Figura 2.15 B**).

No processo de *crossing over*, existem rupturas nas fitas de DNA, que são reparadas de forma que segmentos das cromátides não irmãs sejam trocados (**Figura 2.15 C**). A base molecular deste processo será descrita com mais detalhes no Capítulo 12. O importante é que, após o *crossing over*, as duas cromátides-irmãs não são mais idênticas: uma cromátide tem alelos *A* e *B* enquanto sua cromátide-irmã (a cromátide submetida a *crossing over*) tem os alelos *a* e *B*. Da mesma forma, uma cromátide de outro cromossomo tem os alelos *a* e *b* e a outra cromátide tem os alelos *A* e *b*. Cada uma das quatro cromátides agora tem uma combinação única de alelos: *A B*, *a B*, *A b* e *a b*. Os dois cromossomos homólogos acabam se separando, cada um indo para uma célula diferente. Na meiose II, as duas cromátides de cada cromossomo se separam e cada uma das quatro células resultantes da meiose tem uma combinação diferente de alelos (**Figura 2.15 D**).

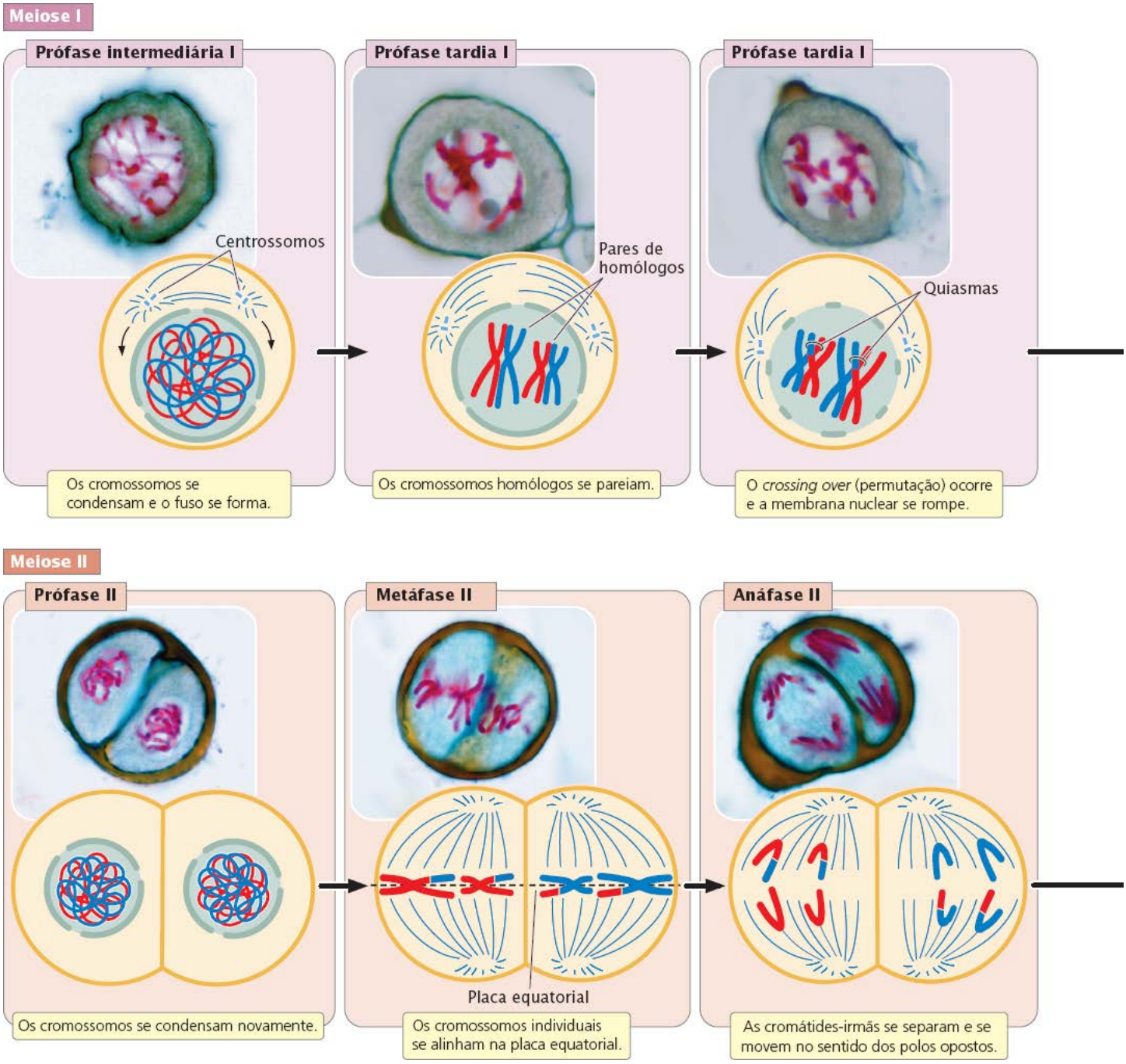
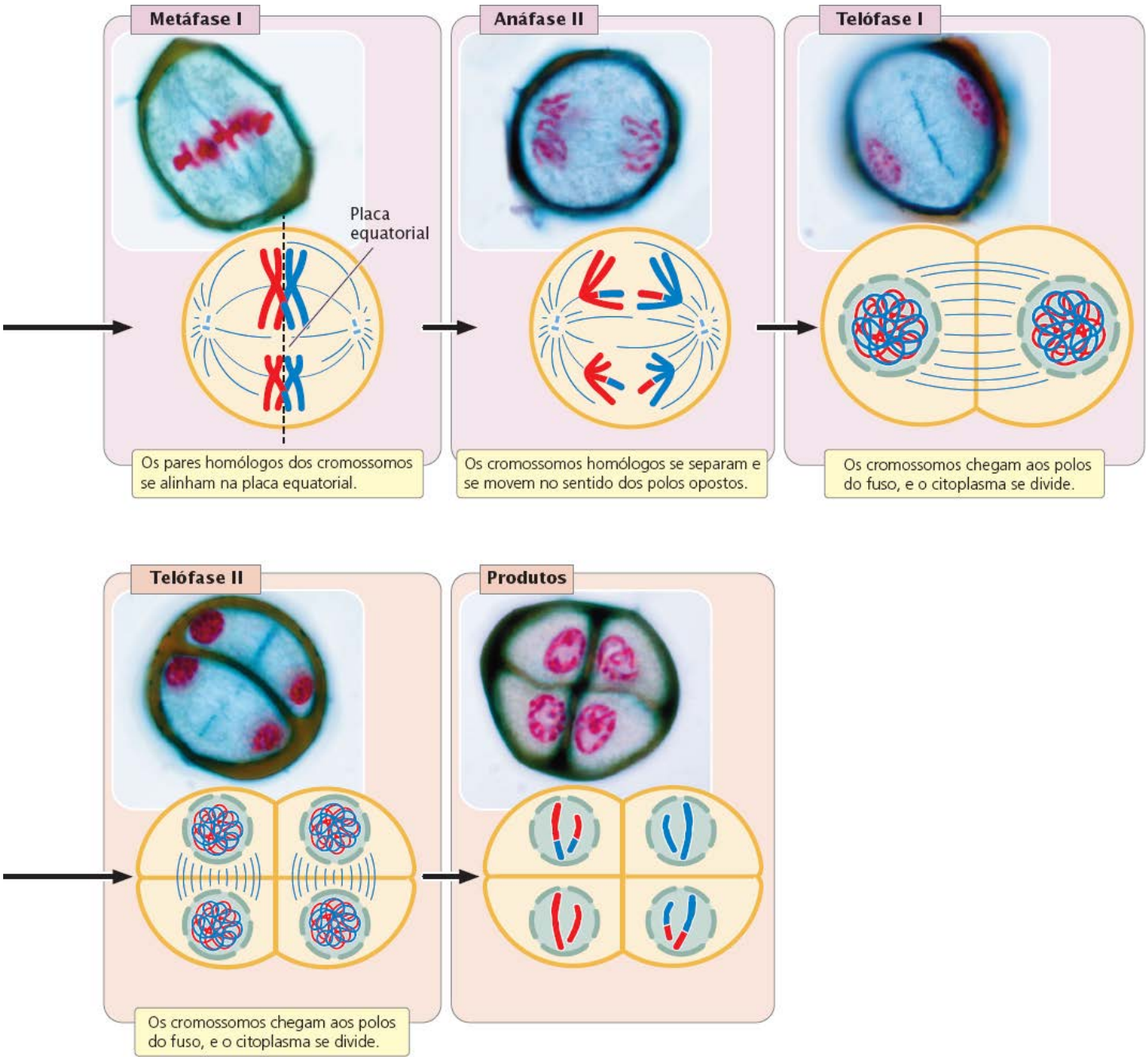


Figura 2.14 A meiose é dividida em estágios. (Fotografias de C. A. Hasenkampf/Biological Photo Service.)

Separação aleatória dos cromossomos homólogos. O segundo processo de meiose que contribui para a variação genética é a distribuição aleatória de cromossomos na anáfase I, após seu alinhamento aleatório na metáfase I. Para ilustrar este processo, considere uma célula com três pares de cromossomos, I, II e III (**Figura 2.16 A**). Um cromossomo de cada par é de origem materna (I_m, II_m e III_m); o outro é de origem paterna (I_p, II_p e III_p). Os pares de cromossomos se alinham no centro da célula na metáfase I, e na anáfase I os cromossomos de cada par homólogo se separam.

Como cada par de homólogos se alinha e se separa é aleatório e independente de como outros pares de cromossomos se alinham e se separam (**Figura 2.16 B**). Por acaso, todos os cromossomos maternos podem migrar para um lado com todos os cromossomos paternos migrando para outro lado. Após a divisão, uma célula teria os cromossomos I_m, II_m e III_m, e a outra I_p, II_p e III_p. Por outro lado, os cromossomos I_m, II_m e III_p podem se mover para um lado, e os cromossomos I_p, II_p e III_m para outro. As diferentes migrações produziriam diferentes combinações de cromossomos nas células resultantes (**Figura 2.16 C**). Existem quatro formas pelas quais uma célula diploide com três pares de cromossomos pode se dividir, produzindo um total de oito combinações diferentes de cromossomos nos gametas. Em geral, o número de possíveis combinações é 2^n, em que n é igual ao número de pares homólogos. À medida que o número de pares de cromossomos aumenta, o número de combinações se torna muito maior. Nos seres humanos, que têm 23 pares de cromossomos, são 2^{23} ou 8.388.608 diferentes combinações de cromossomos possíveis a partir da separação aleatória dos cromossomos homólogos. As consequências genéticas do processo chamado segregação independente serão exploradas com mais detalhes no Capítulo 3.



Em resumo, o *crossing over* desloca os alelos no *mesmo* cromossomo em novas combinações, enquanto a distribuição aleatória dos cromossomos maternos e paternos embaralha os alelos em *diferentes* cromossomos em novas combinações. Juntos, estes dois processos conseguem provocar substancial variação genética entre as células oriundas da meiose. **Resolva os Problemas 33 e 34**

Conceitos

Os dois mecanismos que produzem variação genética na meiose são o *crossing over* (ou permutação) e a distribuição aleatória dos cromossomos maternos e paternos.

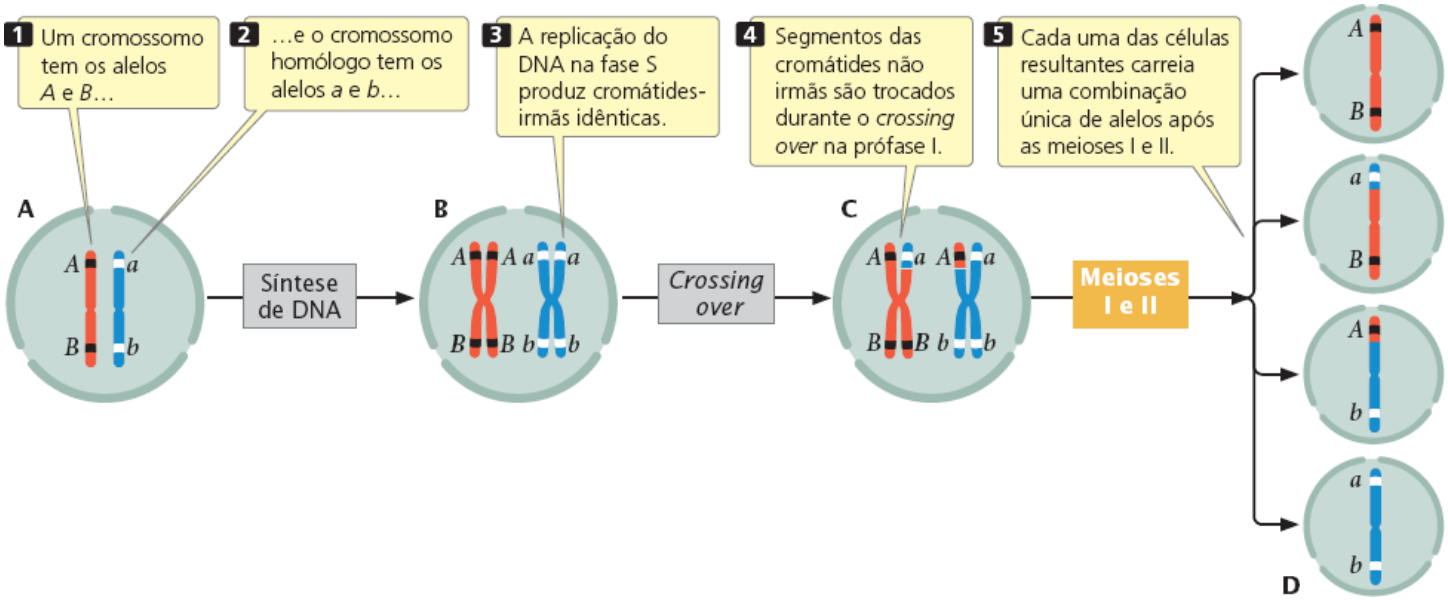


Figura 2.15 **O *crossing over* produz variação genética.**

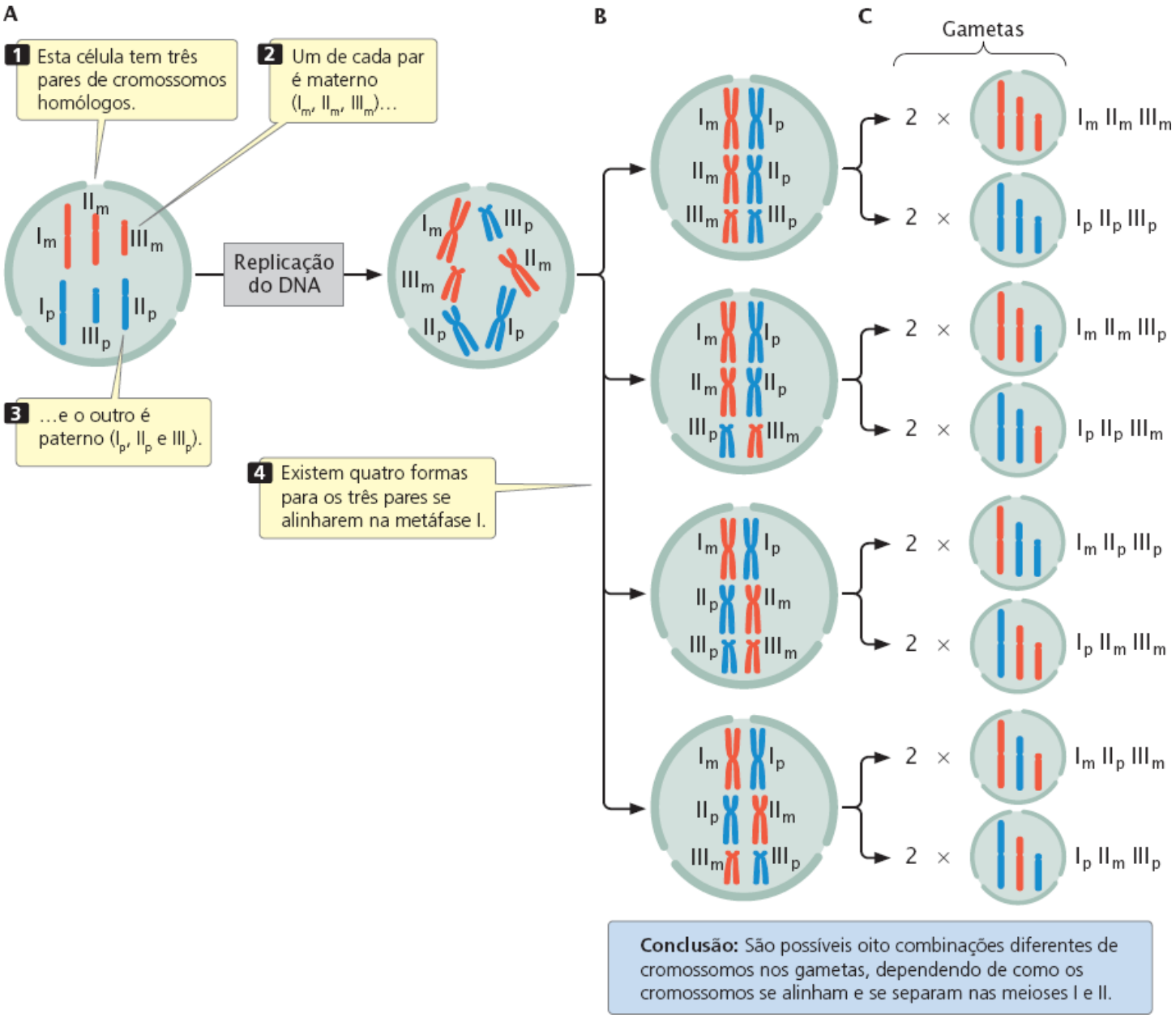


Figura 2.16 **A variação genética é produzida por meio da distribuição aleatória de cromossomos na meiose.** Neste exemplo, a célula tem três pares de cromossomos homólogos.

Conceitos conectantes

Mitose e meiose comparadas

Agora que já examinamos os detalhes da mitose e meiose, vamos comparar os dois processos (**Figura 2.17** e **Quadro 2.3**). Tanto na mitose quanto na meiose, os cromossomos se contraem e se tornam visíveis, existe movimento dos cromossomos para os polos do fuso e são acompanhados pela divisão celular. As semelhanças acabam aqui e a partir deste ponto os processos são bem diferentes.

A mitose resulta em uma única divisão celular e geralmente produz duas células-filhas. A meiose, por outro lado, inclui duas divisões celulares e em geral produz quatro células. Nas células diploides, existem cromossomos homólogos antes da meiose e da mitose, mas o pareamento dos homólogos ocorre apenas na meiose.

Outra diferença é que, na meiose, o número de cromossomos é dividido à metade como consequência da separação dos pares homólogos dos cromossomos na anáfase I, mas não ocorre redução no número de cromossomos na mitose. Além disso, a meiose é caracterizada por dois processos que produzem a variação genética: o *crossing over* (na prófase I) e a distribuição aleatória dos cromossomos maternos e paternos (na anáfase I). Não existem processos equivalentes na mitose.

A mitose e a meiose também diferem no comportamento dos cromossomos na metáfase e anáfase. Na metáfase I da meiose, os *pares homólogos* dos cromossomos se alinham na placa equatorial, enquanto os *cromossomos individuais* se alinham na placa equatorial na metáfase da mitose (e na metáfase II da meiose). Na anáfase I da meiose, os *cromossomos pareados* se separam e migram para os polos opostos do fuso, cada cromossomo com duas cromátides presas ao centrômero. Em contraste, na anáfase da mitose (e na anáfase II da meiose) as *cromátides-irmãs* se separam e cada cromossomo que se move para um polo do fuso não é replicado. **Resolva os Problemas 28 e 29**

Separação das cromátides-irmãs e dos cromossomos homólogos

Recentemente, foram identificadas algumas das moléculas necessárias para unir e separar as cromátides e cromossomos homólogos. A **coesina**, uma proteína que mantém as cromátides juntas, é a chave para o comportamento dos cromossomos na mitose e na meiose (**Figura 2.18 A**). As cromátides-irmãs são mantidas unidas pela coesina, que é estabilizada na fase S e persiste pela fase G_2 e perto da mitose. Na anáfase da mitose, a coesina ao longo da extensão complexa do cromossomo é degradada por uma enzima chamada separase, permitindo que as cromátides-irmãs se separem.

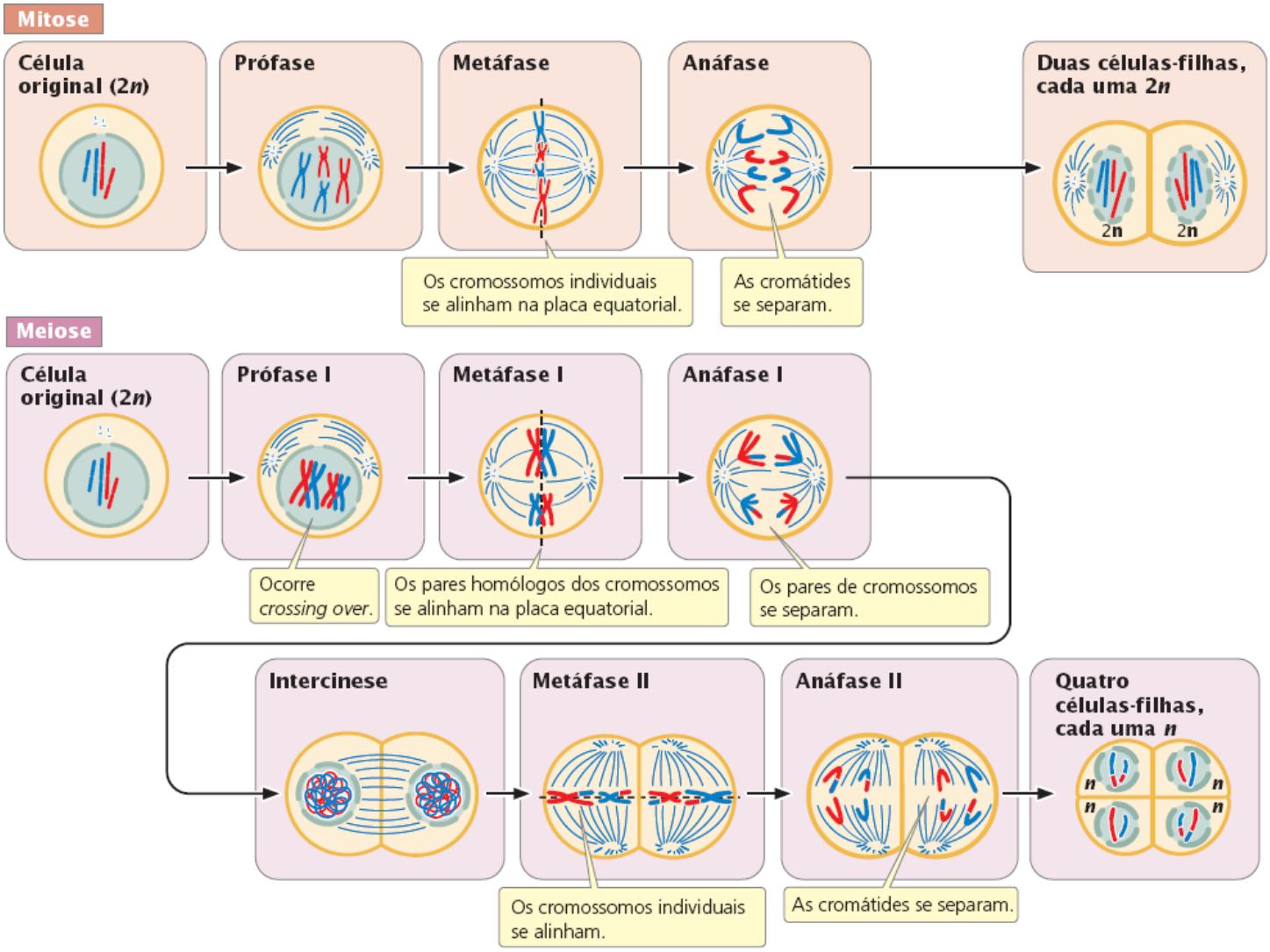


Figura 2.17 **Mitose e meiose comparadas.**

Como observamos, a mitose e meiose diferem fundamentalmente no comportamento dos cromossomos na anáfase (ver Figura 2.17). Por que os homólogos se separam na anáfase I da meiose, enquanto as cromátides se separam na anáfase da mitose e na anáfase II da meiose? É importante observar que as formas de coesina usadas na mitose e meiose são diferentes. No início da meiose, a coesina específica da meiose é encontrada ao longo de toda a extensão dos braços de um cromossomo (**Figura 2.18 B**). A coesina também atua nos braços do cromossomo dos homólogos nos quiasmas, prendendo os dois homólogos nas suas extremidades.

Quadro 2.3 Comparação entre mitose, meiose I e meiose II.

Evento	Mitose	Meiose I	Meiose II
Divisão celular	Sim	Sim	Sim
Redução do cromossomo	Não	Sim	Não
Variação genética produzida	Não	Sim	Não
Crossing over	Não	Sim	Não
Distribuição aleatória dos cromossomos maternos e paternos	Não	Sim	Não
Metáfase	Os cromossomos individuais se alinham	Os pares homólogos se alinham	Os cromossomos individuais se alinham

Anáfase	As cromátides se separam	Cromossomos homólogos se separam	As cromátides se separam

Na anáfase I, a coesina ao longo dos braços do cromossomo é degradada, permitindo que os dois homólogos se separem. Entretanto, a coesina no centrômero é protegida por uma proteína chamada de shugoshina, que significa “espírito guardião” em japonês. Graças a esta ação protetora da shugoshina, a coesina do centrômero permanece intacta e evita a separação das duas cromátides-irmãs durante a anáfase I da meiose. Posteriormente, a shugoshina é degradada. Ao término da metáfase II, a coesina do centrômero – não mais protegida – é degradada, permitindo que as cromátides-irmãs se separem na anáfase II, da mesma forma que na mitose (ver Figura 2.18 B). Resolva o Problema 30

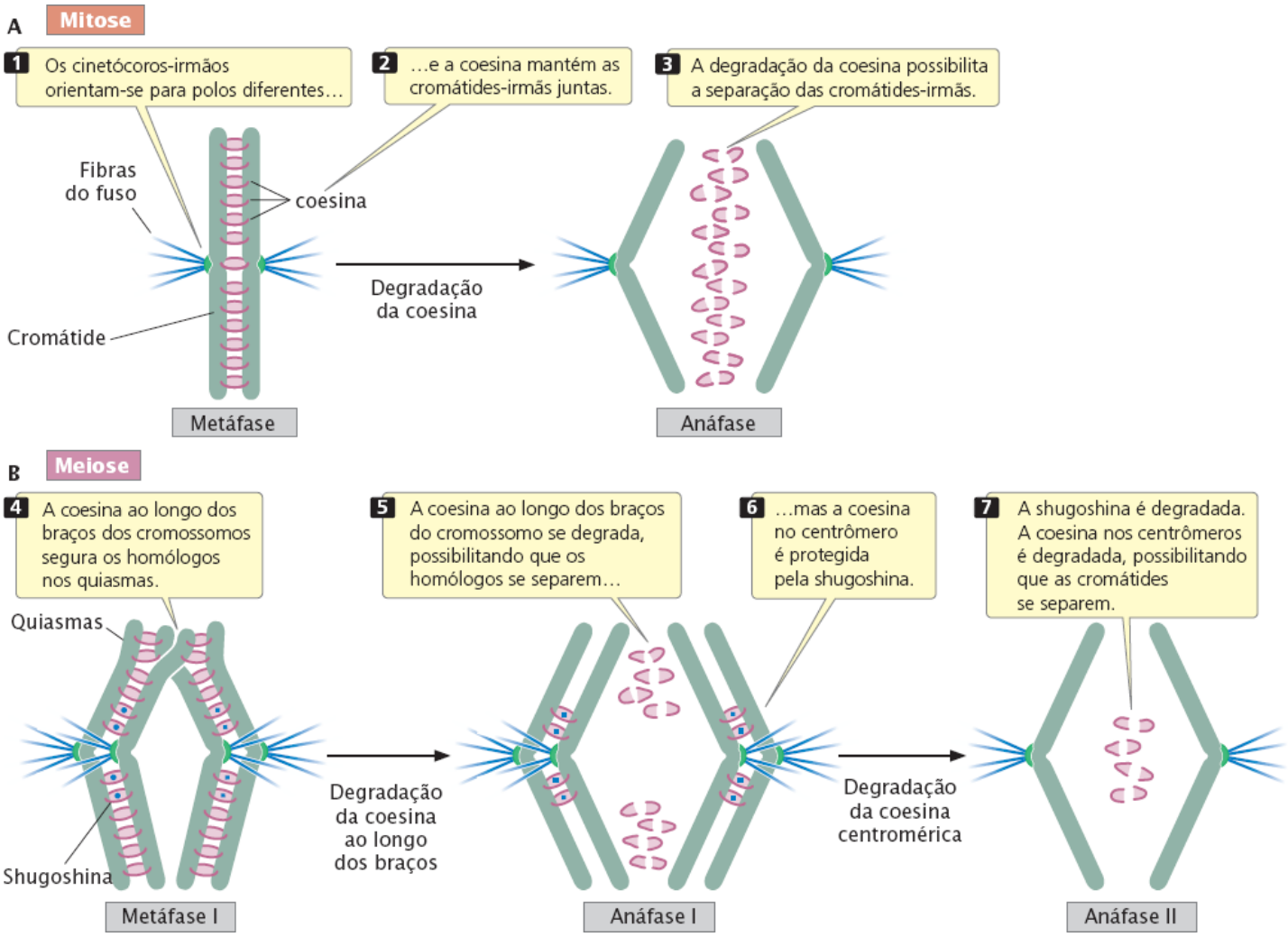


Figura 2.18 A coesina controla a separação das cromátides e cromossomos na mitose e meiose.

Conceitos

A coesina mantém as cromátides-irmãs juntas durante a parte inicial da mitose. Na anáfase, a coesina é degradada, permitindo a separação das cromátides-irmãs. Na meiose, a coesina é protegida nos centrômeros durante a anáfase I e os cromossomos homólogos também, mas não as cromátides-irmãs, que se separam na meiose I. A degradação da coesina do centrômero permite que as cromátides-irmãs se separem na anáfase II da meiose.

Checagem dos conceitos 6

Como a shugoshina afeta as cromátides-irmãs na meiose I e meiose II?

Meiose nos ciclos de vida dos animais e das plantas

O resultado geral da meiose são quatro células haploides geneticamente variáveis. Vamos ver onde a meiose se encaixa nos

ciclos de vida de um animal ou planta multicelular.

Meiose nos animais. A produção de gametas no macho, um processo chamado **espermatogênese**, ocorre nos testículos. Aqui, as células germinativas diploides primordiais se dividem por mitose para produzirem células diploides chamadas de **espermatogônias** (**Figura 2.19 A**). Cada espermatogônia sofre ciclos repetidos de mitose, dando origem a numerosas espermatogônias adicionais. Alternativamente, uma espermatogônia pode iniciar a meiose e entrar na prófase I. Agora chamada de **espermatócito primário**, a célula ainda é diploide porque os cromossomos homólogos não foram separados ainda. Cada espermatócito primário completa a meiose I, dando origem a dois **espermatócitos secundários** haploides que então sofrem meiose II com cada um produzindo duas **espermátides** haploides. Assim, cada espermatócito primário produz um total de quatro espermátides haploides, que amadurecem e se desenvolvem em espermatozoides.

A produção de gametas nas fêmeas, um processo chamado de **ovogênese (oogênese)**, inicia-se da mesma forma que a espermatogênese. Dentro dos ovários, as células germinativas primordiais diploides se dividem por mitose para produzir **ovogônias** (**Figura 2.19 B**). Como as espermatogônias, as ovogônias (oogônias) podem sofrer ciclos repetidos de mitose ou entrar na meiose. Quando elas entram na prófase, estas células ainda diploides são chamadas de **ovócitos (oócitos) primários**. Cada ovócito (oócito) primário completa a meiose I e se divide.

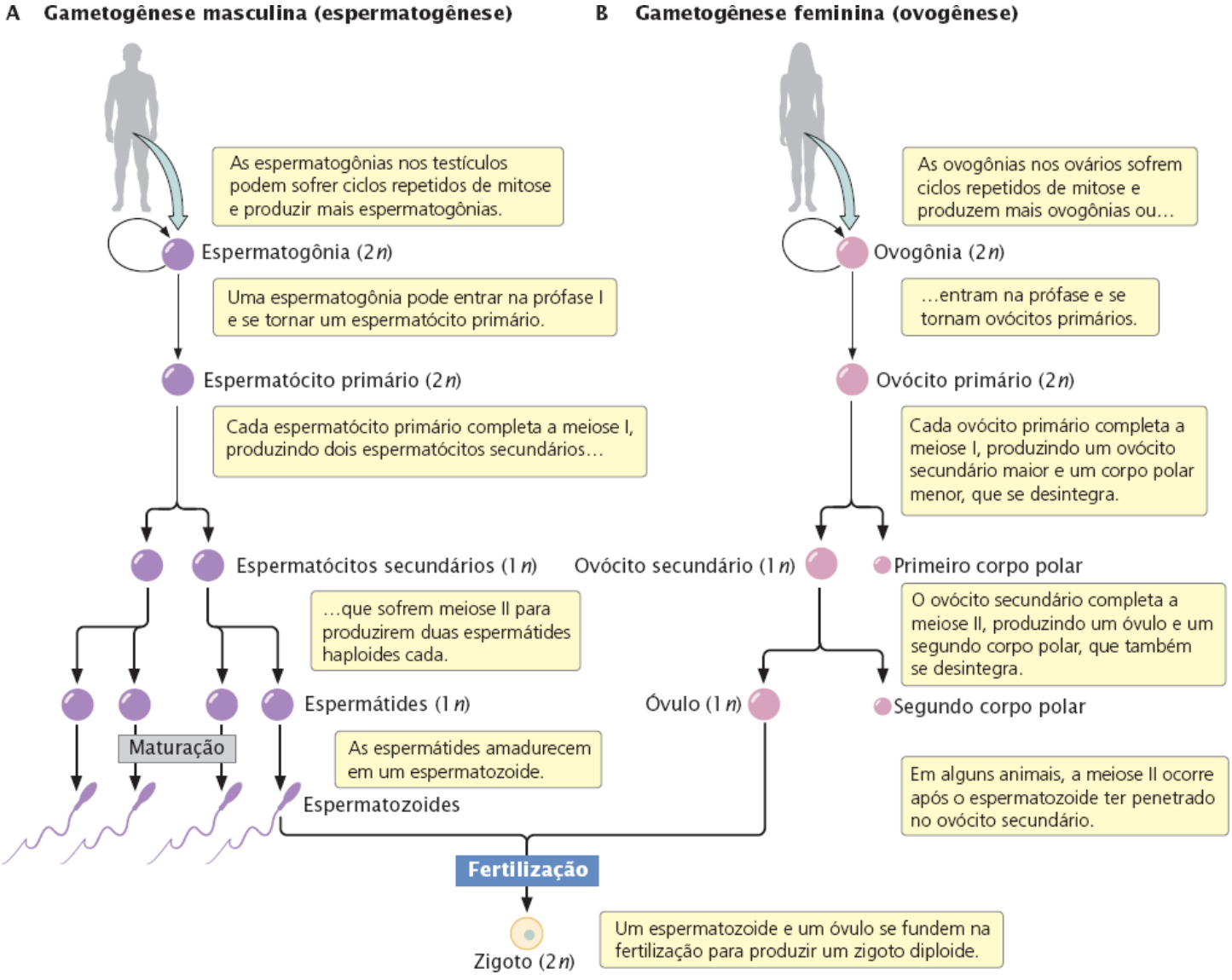


Figura 2.19 Formação de gametas em animais.

Neste ponto, o processo de ovogênese começa a se diferenciar do processo de espermatogênese. Na ovogênese, a citocinese não é equilibrada: a maior parte do citoplasma é alocada em uma das duas células haploides, o **ovócito (oócito) secundário**. A célula menor, que tem metade dos cromossomos, mas apenas uma pequena parte do citoplasma, é chamada de **primeiro corpo polar**; ela pode se dividir ou não. O ovócito secundário completa a meiose e novamente, a citocinese não é equilibrada, com a maior parte do citoplasma passando para uma das células. A célula maior, que adquire a maior parte do citoplasma, é o **óvulo**, o gameta feminino maduro. A célula menor é o **segundo corpo polar**. Apenas o óvulo é capaz de ser fertilizado e os corpos

polares em geral se desintegram. A ovogênese, então, produz um único gameta maduro a partir de cada ovócito primário.

Nos mamíferos, a ovogênese é diferente da espermatogênese de outra forma. A formação do espermatozoide é contínua nos machos durante toda sua vida reprodutiva. Entretanto, a formação dos gametas femininos é um processo descontínuo e ocorre durante alguns anos. A ovogênese começa antes do parto; neste momento, as ovogônias iniciam a meiose e dão origem aos ovócitos primários. A meiose é interrompida e para na prófase I. Uma fêmea nasce com os ovócitos primários parados na prófase I. Nos seres humanos, este período de animação suspensa pode durar 30 ou 40 anos. Antes da ovulação, a elevação dos níveis hormonais estimula um ou mais dos ovócitos primários a recomeçar a meiose. A primeira divisão da meiose é encerrada e um ovócito secundário ovula a partir do ovário. Nos seres humanos e em muitas outras espécies, a segunda divisão da meiose é então adiada até entrar em contato com o espermatozoide. Quando o espermatozoide penetra a camada mais externa do ovócito secundário, a segunda divisão meiótica ocorre, o segundo corpo polar é expulso do ovo e os núcleos do espermatozoide e do óvulo recém-formado se fundem, dando origem ao zigoto.

Conceitos

Nos testículos, uma espermatogônia diploide sofre meiose, produzindo um total de quatro espermatozoides haploides. No ovário, uma ovogônia diploide sofre meiose para produzir um único óvulo grande e corpos polares menores que normalmente se desintegram.

Checagem dos conceitos 7

Um espermatócito secundário tem 12 cromossomos. Quantos cromossomos serão encontrados no espermatócito primário que deu origem ao espermatócito secundário?

a. 6.
b. 12.
c. 18.
d. 24.

Meiose nas plantas. A maioria das plantas tem um ciclo de vida complexo que inclui duas estruturas distintas (gerações): um esporófito diploide secundário e um gametófito haploide multicelular. Estas duas gerações se alternam, o esporófito produz esporos haploides por meio da meiose e o gametófito produz gametas haploides pela mitose (**Figura 2.20**). Este tipo de ciclo de vida é chamado de alternância de gerações. Nele, os produtos imediatos da meiose são chamados de esporos, não gametas; os esporos sofrem uma ou mais divisões mitóticas para produzir gametas. Embora os termos usados para este processo sejam diferentes dos comumente usados para os animais (e de alguns dos empregados até o momento neste capítulo), os processos nas plantas e animais são os mesmos: em ambos a meiose leva à redução no número de cromossomos, produzindo células haploides.

Nas plantas com flores, o esporófito é a parte vegetativa da planta, o gametófito se resume a algumas células haploides dentro do esporófito. A flor, que é parte do esporófito, possui as estruturas reprodutoras. Em algumas plantas, as estruturas reprodutoras masculina e feminina são encontradas na mesma flor, em outras plantas, elas estão em flores diferentes. Em ambos os casos, a parte masculina da flor, o estame, tem as células reprodutoras diploides, chamadas de **microesporócitos**, e cada uma sofre meiose para produzir quatro **microesporos** haploides (**Figura 2.21 A**). Cada microesporo se divide por mitose, produzindo um grão de pólen imaturo com dois núcleos haploides. Um destes núcleos, chamado de núcleo tubo, direciona o crescimento de um tubo de pólen. O outro, chamado de núcleo germinativo, sofre divisão mitótica para produzir duas sementes. O grão de pólen, com seus núcleos haploides, é o gametófito masculino.

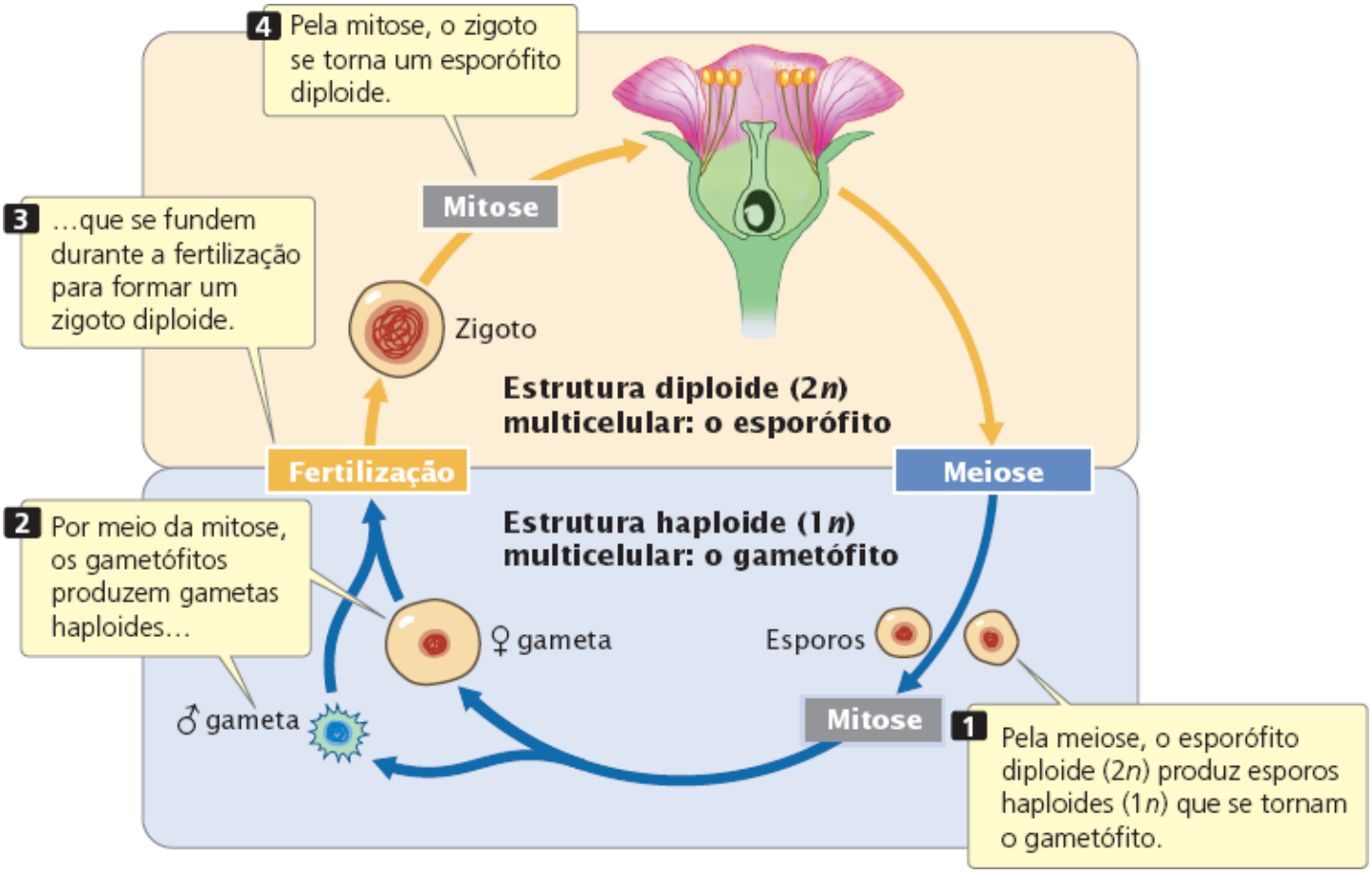


Figura 2.20 **As plantas alternam entre estágios diploide e haploide de vida (feminina, ♀; masculina, ♂).**

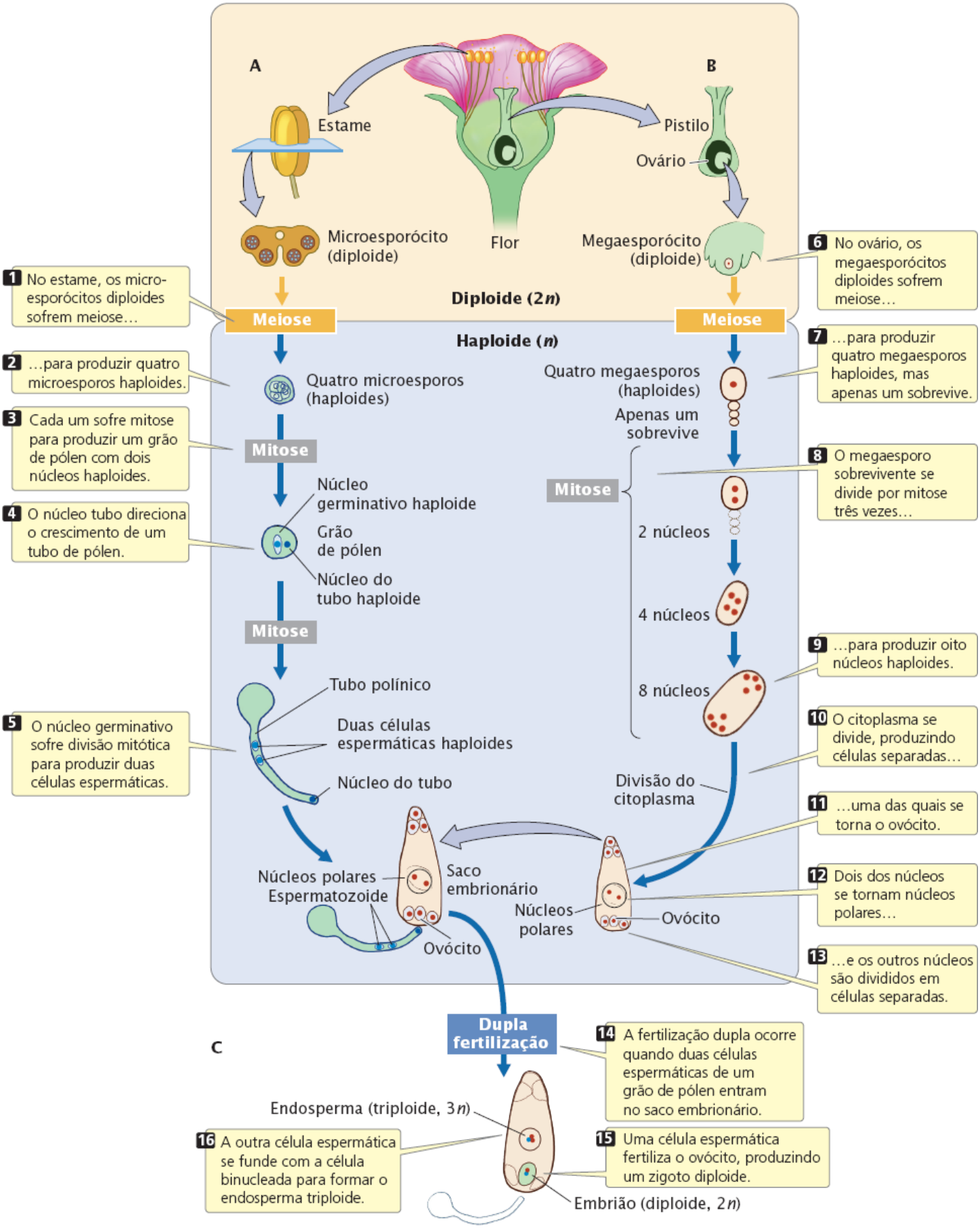


Figura 2.21 Reprodução sexuada nas plantas com flores.

A parte feminina da flor, o ovário, tem células diploides chamadas de **megaesporócitos**; cada uma sofre meiose para produzir quatro **megaesporos** haploides (**Figura 2.21 B**) e apenas uma delas sobrevive. O núcleo do megaesporo sobrevivente se divide por mitose três vezes, produzindo um total de oito núcleos haploides que compõem o gametófito feminino, também conhecido como saco embrionário. A divisão do citoplasma produz células separadas, uma das quais se torna o ovócito (oócito).

Quando a planta floresce, os estames se abrem e liberam os grãos de pólen. O pólen pousa no estigma de uma flor – uma

plataforma pegajosa que fica no topo do talo longo chamado de estilo. O ovário fica na base do estilo. Se um grão de pólen germina, ele cresce para baixo para o ovário. As duas células espermáticas passam por esse tubo e entram no saco embrionário (**Figura 2.21 C**). Uma das células espermáticas fertiliza o ovócito (oócito), produzindo um zigoto diploide, que se desenvolve em embrião. A outra célula espermática se funde com os dois núcleos presentes em uma única célula, dando origem a um endosperma 3*n* (triploide), que armazena nutrientes que serão usados posteriormente pela planta embrionária. Estes dois eventos de fertilização são chamados de fertilização dupla.

Conceitos

No estame de uma planta com flores, a meiose produz microesporos haploides que se dividem por mitose para produzirem uma célula espermática haploide no grão de pólen. No ovário, a meiose produz quatro megaesporos haploides e apenas um destes se divide por mitose três vezes para produzir oito núcleos haploides. Após a polinização, uma célula espermática fertiliza o ovócito (oócito), produzindo um zigoto diploide; a outra se funde com dois núcleos para formar o endosperma.

Checagem dos conceitos 8

Qual destas estruturas é diploide?

a. Microesporo.
b. Ovócito (oócito).
c. Megaesporo.
d. Microesporócito.

Agora examinamos a ação da meiose no ciclo sexual de dois organismos, um típico animal multicelular e uma planta com flores. Esses ciclos são apenas dois de muitas variações encontradas entre os organismos eucarióticos. Embora os eventos celulares que produzam as células reprodutoras em plantas e animais sejam diferentes no número de divisões celulares, no número de gametas haploides produzidos e no tamanho relativo dos produtos finais, o resultado geral é o mesmo: a meiose dá origem a células haploides, geneticamente variáveis que então se fundem durante a fertilização para produzir uma prole diploide. **Resolva os Problemas 36 e 38**

Resumo dos conceitos

- Uma célula procariótica tem uma estrutura simples, sem envelope nuclear e em geral um único cromossomo circular. Uma célula eucariótica tem estrutura mais complexa, com um núcleo e múltiplos cromossomos lineares que consistem em DNA complexado a histonas
- A reprodução celular requer a cópia do material genético, separação das cópias e divisão celular
- Em uma célula procariótica, o único cromossomo se replica, cada cópia se move no sentido dos lados opostos da célula e a célula se divide. Nas células eucarióticas, a reprodução é mais complexa que nas células procarióticas, exigindo mitose e meiose para garantir que um conjunto completo de informações genéticas seja transferido para cada célula nova
- Nas células eucarióticas, os cromossomos são tipicamente encontrados em pares homólogos. Cada cromossomo funcional consiste em um centrômero, telômeros e múltiplas origens de replicação. Após um cromossomo ser copiado, as duas cópias permanecem presas pelo centrômero, formando cromátides-irmãs
- O ciclo celular consiste em estágios por meio dos quais uma célula eucariótica passa entre as divisões celulares. Ele apresenta (1) interfase, na qual a célula cresce e se prepara para divisão e (2) fase M, na qual ocorrem as divisões nuclear e celular. A fase M consiste em (1) mitose, o processo de divisão nuclear e (2) citocinese, a divisão do citoplasma
- A progressão pelo ciclo celular é controlada por pontos de verificação que regulam o ciclo celular ao permitir ou proibir a célula de seguir para o próximo estágio
- Geralmente a mitose leva à produção de duas células geneticamente idênticas
- A reprodução sexuada produz proles geneticamente variáveis e permite acelerar a evolução. Isto inclui a meiose, na qual as células sexuais haploides são produzidas, e fertilização, a fusão das células sexuais. A meiose inclui duas divisões celulares. Na meiose I, ocorre *crossing over* (ou permutação) e os cromossomos homólogos se separam. Na meiose II, as cromátides se separam
- O resultado habitual da meiose é a produção de quatro células haploides geneticamente variáveis. A variação genética na meiose é produzida pelo *crossing over* e pela distribuição aleatória de cromossomos maternos e paternos
- A coesina mantém as cromátides-irmãs unidas. Na metáfase da mitose e na metáfase II da meiose, a degradação da coesina

permite que as cromátides-irmãs se separem. Na meiose I, a coesina centromérica permanece intacta e mantém as cromátides-irmãs juntas de modo que os cromossomos homólogos, mas não as cromátides-irmãs, separem-se na anáfase I

- Nos animais, uma espermatogônia diploide sofre meiose, produzindo quatro espermatozoides haploides. Uma ovogônia (oogônia) diploide sofre meiose para produzir um óvulo haploide grande e um ou mais corpos polares menores
- Nas plantas, um microesporócito diploide no estame sofre meiose para produzir quatro grãos de pólen, cada um com duas células espermáticas haploides. No ovário, um megaesporócito diploide sofre meiose para produzir oito núcleos haploides, um dos quais forma o ovócito (oócito). Durante a polinização, uma célula espermática fertiliza o ovócito (oócito), produzindo um zigoto diploide, a outra se funde com dois núcleos para formar o endosperma 3*n*.

Termos importantes

Anáfase
Anáfase I
Anáfase II
Archaea
Bivalente
Ciclo celular
Citocinese
Coesina
Cromátide-irmã
Cromatina
Crossing over
Diploide
Espermátide
Espermatócito primário
Espermatócito secundário
Espermatogênese
Espermatogônia
Eubactérias
Eucarioto
Fase M (mitótica)
Fertilização
Haploide
Histona
Intercinese
Interfase
Megaesporo
Megaesporócito
Meiose
Metáfase
Metáfase I
Metáfase II
Microesporo
Microesporócito
Mitose
Núcleo
Origem de replicação
Ovócito (oócito) primário
Ovócito (oócito) secundário
Ovogênese (oogênese)
Ovogônia (oogônia)
Óvulo
Par homólogo
Poliploide
Ponto de verificação

Primeiro corpo polar
Procarioto
Prófase
Prófase I
Prófase II
Pró-metáfase
Recombinação
Segundo corpo polar
Sinapse
Telófase
Telófase I
Telófase II
Telômero
Tétrade
Vírus

Respostas da Checagem dos conceitos

1. Eubactérias e archaea são procariotos. Eles são diferentes dos eucariotos porque não têm núcleo, seu genoma consiste em um único cromossomo circular e uma pequena quantidade de DNA.
2. b.
3. O cinetócoro é o ponto no qual os microtúbulos do fuso se prendem ao cromossomo. Se não houver cinetócoros, os microtúbulos do fuso não se prendem ao cromossomo, o cromossomo não será atraído para o núcleo e as células resultantes terão um cromossomo a menos.
4. a.
5. c.
6. Durante a anáfase I, a shugoshina protege a coesina nos centrômeros da ação da separase; então a coesina permanece intacta e as cromátides-irmãs permanecem unidas. Em seguida, a shugoshina é degradada, e a coesina do centrômero é clivada na anáfase II e as cromátides se separam.
7. d.
8. d.

Problemas desenvolvidos

Problema 1

Um estudante examina uma secção fina de uma ponta de raiz de cebola e registra o número de células que estão em cada estágio do ciclo celular. Ela observa 94 células na interfase, 14 células na prófase, 3 células na pró-métafase, 3 células na metáfase, 5 células na anáfase e 1 célula na telófase. Se o ciclo celular completo em uma ponta da raiz da cebola leva 22 h, qual é a duração média de cada estágio do ciclo? Considere que todas as células estão no ciclo celular ativo (não G_0).

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
A duração média de cada estágio do ciclo celular.

Quais informações para solucionar o problema são fornecidas?

- O número de células em diferentes estágios do ciclo celular.
- Um ciclo completo requer 22 h.

Para a solução deste problema, revise:
O ciclo celular e a mitose, na Seção 2.2.

Etapas para a solução

Este problema é solucionado em duas etapas. Primeiro, calculamos as proporções de células em cada

estágio do ciclo celular, que correspondem ao intervalo de tempo que uma célula comum leva em cada estágio. Por exemplo, se as células consomem 90% do seu tempo na interfase, então, em um dado momento, 90% das células estará na interfase. A segunda etapa é converter as proporções em períodos de tempo, o que é feito ao multiplicar as proporções pelo período total do ciclo celular (22 h).

ETAPA 1 | Calcular a proporção de células em cada estágio
A proporção de células em cada estágio é igual ao número de células encontradas naquele estágio, dividido pelo número total de células examinadas:

Interfase	94/120 = 0,783
Prófase	14/120 = 0,117
Pró-metáfase	3/120 = 0,025
Metáfase	3/120 = 0,025
Anáfase	5/120 = 0,042
Telófase	1/120 = 0,008

Dica: O total d e todas as proporções deve ser 1,0.

ETAPA 2 | Determinar a duração média de cada estágio
Para determinar a duração média de cada estágio, multiplique a proporção de células em cada estágio pelo tempo necessário para o ciclo celular completo:

Interfase	0,783 × 22 h = 17,23 h
Prófase	0,117 × 22 h = 2,57 h
Pró-metáfase	0,025 × 22 h = 0,55 h
Metáfase	0,025 × 22 h = 0,55 h
Anáfase	0,042 × 22 h = 0,92 h
Telófase	0,008 × 22 h = 0,18 h

Dica: O tempo total para todos os estágios deve ser 22 h.

Problema 2

Uma célula em G_1 da interfase tem 8 cromossomos. Quantos cromossomos e moléculas de DNA serão encontrados por célula à medida que ela passa pelos seguintes estágios: G_2, metáfase da mitose, anáfase da mitose, após citocinese na mitose, metáfase I da meiose, metáfase II da meiose e após citocinese da meiose II?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
O número de cromossomos e o número de moléculas de DNA por célula em diferentes estágios do ciclo celular e meiose.

Quais informações para solucionar o problema são fornecidas?

- Uma célula em G_1 tem 8 cromossomos
- Diferentes estágios do ciclo celular e meiose.

Para a solução deste problema, revise:

Conceitos conectantes | *Contagem de cromossomos e moléculas de DNA*, na Seção 2.2.

Etapas para a solução

Lembre as regras sobre a contagem de cromossomos e moléculas de DNA: (1) para determinar o número de cromossomos, contar o número de centrômeros funcionais; (2) para determinar o número de moléculas de DNA, determinar se existem cromátides-irmãs. Se existirem cromátides-irmãs, o número de moléculas de DNA é o dobro do número de cromossomos. Se os cromossomos não forem replicados (não têm cromátides-irmãs), o número de moléculas de DNA é igual ao número de cromossomos. Pense bem sobre quando e como o número de cromossomos e de moléculas de DNA alteram no curso da mitose e meiose.

Dica: Essas duas regras são importantes para responder a pergunta.

O número de moléculas de DNA aumenta apenas na fase S, quando o DNA é replicado, o número de moléculas de DNA reduz apenas quando a célula se divide.

O número de cromossomos aumenta apenas quando as cromátides-irmãs se separam na anáfase da mitose e na anáfase II da meiose (os cromossomos homólogos, não as cromátides, se separam na anáfase I da meiose). Assim como o número de moléculas de DNA, o número de cromossomos também é reduzido apenas pela divisão celular.

Vamos aplicar estes princípios no problema. Uma célula em G_1 tem 8 cromossomos e não há cromátides-irmãs; então existem 8 moléculas de DNA em G_1. O DNA se replica na fase S e agora cada cromossomo consiste em duas cromátides; então em G_2, $2 \times 8 = 16$ moléculas de DNA em cada célula. Entretanto, as duas cópias de cada molécula de DNA permanecem presas ao centrômero, então ainda existem apenas 8 cromossomos. À medida que a célula passa pela prófase e metáfase do ciclo celular, o número de cromossomos e o número de moléculas de DNA permanecem os mesmos; então, na metáfase, existem 16 moléculas de DNA e 8 cromossomos. Na anáfase, as cromátides se separam e cada uma se torna um cromossomo independente; neste momento, o número de cromossomos aumenta de 8 para 16. Este aumento é temporário, até a célula se dividir na telófase ou após esta fase. O número de moléculas de DNA permanece 16 na anáfase. O número de moléculas de DNA e cromossomos por célula é reduzido por citocinese após a telófase, porque os 16 cromossomos e moléculas de DNA agora estão distribuídos entre duas células. Portanto, após a citocinese, cada célula tem 8 moléculas de DNA e 8 cromossomos, os mesmos números presentes no início do ciclo celular.

Lembrete: O número de cromossomos aumenta apenas quando as cromátides se separam. O número de moléculas de DNA aumenta apenas na fase S.

Agora, vamos rastrear o número de moléculas de DNA e cromossomos durante a meiose. Em G_1 temos 8 cromossomos e 8 moléculas de DNA. O número de moléculas de DNA aumenta para 16 na fase S, mas o número de cromossomos permanece 8 (cada cromossomo tem duas cromátides). A célula entra na metáfase I com 16 moléculas de DNA e 8 cromossomos. Na anáfase I da meiose, os cromossomos homólogos se separam, mas o número de cromossomos permanece 8. Após a citocinese, os 8 cromossomos originais são distribuídos entre as duas células; o número de cromossomos cai para 4 (cada um com duas cromátides). As 16 moléculas de DNA também são distribuídas entre as duas células; então o número de moléculas de DNA por célula é 8. Não ocorre síntese de DNA na intercinese e cada célula ainda mantém 4 cromossomos e 8 moléculas de DNA durante a metáfase II. Na anáfase II, as duas cromátides de cada cromossomo se separam, aumentando temporariamente o número de cromossomos em cada célula para oito, enquanto o número de moléculas de DNA em cada célula permanece 8. Após a citocinese, os cromossomos e moléculas de DNA são mais uma vez distribuídos entre duas células, fornecendo 4 cromossomos e 4 moléculas de DNA para cada uma. Os resultados estão resumidos no quadro abaixo:

Estágio	**Número de cromossomos por célula**	**Número de moléculas de DNA por célula**
G_1	8	8
G_2	8	16

Metáfase da mitose	8	16
Anáfase da mitose	16	16
Após a citocinese da mitose	8	8
Metáfase I da meiose	8	16
Metáfase II da meiose	4	8
Após a citocinese da meiose II	4	4

Questões de compreensão

Seção 2.1

1. Quais são as diferenças genéticas entre as células procarióticas e eucarióticas?
2. Por que os vírus que infectam as células dos mamíferos são úteis para o estudo da genética dos mamíferos?

Seção 2.2

3. Liste os três eventos fundamentais que têm de ocorrer na reprodução celular.
4. Descreva o processo de reprodução da célula procariótica.
5. Cite três elementos estruturais essenciais de um cromossomo eucariótico funcional e descreva suas funções.
6. Descreva e identifique quatro tipos diferentes de cromossomos com base na posição do centrômero.
7. Liste os estágios da interfase e os principais eventos que ocorrem em cada estágio.
8. O que são pontos de verificação? Liste alguns dos pontos de verificação importantes no ciclo celular.
9. Liste os estágios da mitose e os principais eventos que ocorrem em cada estágio.
10. Descreva resumidamente como os cromossomos se movem para os polos do fuso durante a anáfase.
11. Quais são os resultados importantes do ponto de vista genético do ciclo celular e da mitose?
12. Por que as duas células produzidas pelo ciclo celular são geneticamente idênticas?

Seção 2.3

13. Quais são os estágios da meiose e os principais eventos que ocorrem em cada estágio?
14. Quais são os principais resultados da meiose?
15. Quais dois processos únicos da meiose são responsáveis pela variação genética? Em que ponto na meiose estes processos ocorrem?
16. Liste algumas semelhanças e diferenças entre mitose e meiose. Quais diferenças você acha que são mais importantes e por quê?
17. Explique resumidamente por que cromátides-irmãs permanecem juntas na anáfase I, mas se separam na anáfase II da meiose.
18. Descreva os processos de espermatogênese e ovogênese nos animais.
19. Descreva os processos da formação do gameta masculino e gameta feminino nas plantas.

Questões e problemas aplicados

Introdução

*20. Responda as perguntas a seguir sobre o Enigma dos cegos, apresentado no início do capítulo.
   a. O que as duas meias de um par representam no ciclo celular?
   b. Na charada apresentada, cada cego compra seu próprio par de meias, mas o vendedor coloca todos os pares em uma bolsa. Assim, existem dois pares de meias de cada cor na bolsa (dois pares negros, dois pares azuis, dois pares cinza etc.). O que os dois pares (quatro meias no total) de cada cor representam?
   c. No ciclo celular, qual é o filamento que conecta as duas meias de um par?
   d. No ciclo celular, qual é a faca molecular que corta o filamento que segura as duas meias de um par?
   e. Que item no enigma tem a mesma função das fibras do fuso?
   f. O que aconteceria se um homem não conseguisse agarrar sua meia de um par em particular e como isto tem relação com

os eventos no ciclo celular?

Seção 2.1

21. Uma célula tem um cromossomo circular e não tem membrana nuclear. Seu DNA é complexado a algumas histonas. Esta célula pertence a uma eubactéria, archaea ou eucarioto? Explique seu raciocínio.

Seção 2.2

22. Uma espécie específica tem três pares de cromossomos: um par acrocêntrico, um par metacêntrico e um par submetacêntrico. Desenhe uma célula desta espécie da forma como ela apareceria na metáfase da mitose.

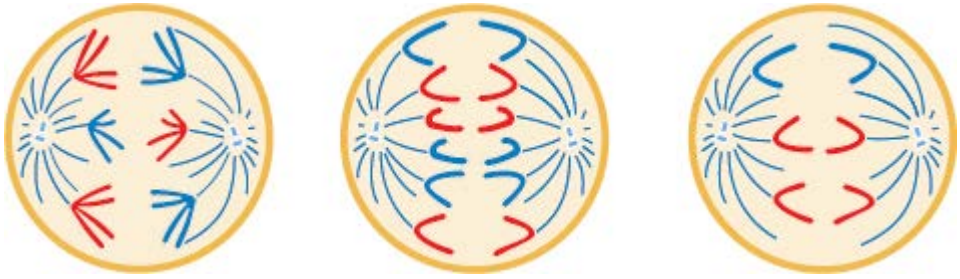
*23. Um biólogo observa um conjunto de células e conta 160 células na interfase, 20 células na prófase, 6 células na pró-metáfase, 2 células na metáfase, 7 células na anáfase e 5 células na telófase. Se o ciclo celular completo leva 24 h, qual é a duração média da fase M nestas células? E da metáfase?

Seção 2.3

24. Uma espécie específica tem três pares de cromossomos: um par acrocêntrico, e dois pares metacêntricos. Desenhe uma célula desta espécie como ela apareceria nos seguintes estágios da meiose.
    a. Metáfase I.
    b. Anáfase I.
    c. Metáfase II.
    d. Anáfase II.
25. Construa um quadro semelhante ao da **Figura 2.11** para os diferentes estágios da meiose, considerando o número de cromossomos por célula e o número de moléculas de DNA por célula para uma célula que inicia com 4 cromossomos (dois pares homólogos) em G_2. Inclua os seguintes estágios no seu quadro: G_1, S, G_2, prófase I, metáfase I, anáfase I, telófase I (após a citocinese), prófase II, metáfase II, anáfase II e telófase II (após citocinese).

*26. Uma célula em G_1 da interfase tem 12 cromossomos. Quantos cromossomos e moléculas de DNA serão encontrados por célula à medida que esta célula original avança nos seguintes estágios:
    a. G_2 da interfase.
    b. Metáfase I da meiose.
    c. Prófase da mitose.
    d. Anáfase I da meiose.
    e. Anáfase II da meiose.
    f. Prófase II da meiose.
    g. Após a citocinese que acompanha a mitose.
    h. Após a citocinese que acompanha a meiose II.
27. Quais são as semelhanças entre os eventos que ocorrem na espermatogênese e ovogênese? E as diferenças?

*28. Todas as células a seguir, apresentadas em vários estágios de mitose e meiose, surgem da mesma espécie rara de planta.
    a. Qual é o número diploide de cromossomos nesta planta?
    b. Dê os nomes de cada estágio de mitose ou meiose apresentados.
    c. Dê o número de cromossomos e moléculas de DNA por célula encontrados em cada estágio.



*29. A quantidade de DNA por célula de uma espécie específica é medida nas células encontradas em vários estágios da meiose e foram obtidas as seguintes quantidades:

Quantidade de DNA por célula

_____ 3,7 pg _____ 7,3 pg _____ 14,6 pg

Combine as quantidades de DNA anterior com os estágios correspondentes do ciclo celular (de *a* a *f*). Você pode usar mais de

um estágio para cada quantidade de DNA.

Estágio da meiose

a. G_1.
b. Prófase I.
c. G_2.
d. Após telófase II e citocinese.
e. Anáfase I.
f. Metáfase II.

*30. Como cada um dos eventos a seguir afetaria o desfecho da mitose ou meiose?
a. A coesina mitótica não é formada no estágio inicial da mitose.
b. A shugoshina está ausente durante a meiose.
c. A shugoshina não se degrada após a anáfase I da meiose.
d. A separase é defeituosa.

*31. Uma célula na prófase II da meiose tem 12 cromossomos. Quantos cromossomos seriam encontrados em uma célula do mesmo organismo se ele estivesse na prófase da mitose? Prófase I da meiose?

32. Uma célula tem 8 cromossomos em G_1 da interfase. Faça um desenho desta célula com seus cromossomos nos seguintes estágios e indique quantas moléculas de DNA são encontradas em cada um deles.
a. Metáfase da mitose.
b. Anáfase da mitose.
c. Anáfase II da meiose.
d. Diplóteno da meiose I.

*33. A mosca-da-fruta *Drosophila melanogaster* (esquerda) tem quatro pares de cromossomos, enquanto a mosca-doméstica *Musca domestica* (direita) tem seis pares de cromossomos. Em qual espécie você esperaria ver mais variações genéticas na prole de um cruzamento? Explique sua resposta.

(Hermann Eisenbeiss; Materz Mali/iStockphoto.)

*34. Uma célula tem dois pares de cromossomos submetacêntricos, que chamaremos de cromossomos I_a, I_b, II_a e II_b (os cromossomos I_a e I_b são homólogos e os cromossomos II_a e II_b são homólogos). O alelo *M* está localizado no braço longo do cromossomo I_a e o alelo *m* está localizado na mesma posição no cromossomo I_b. O alelo *P* está localizado no braço curto do cromossomo I_a e o alelo *p* está localizado na mesma posição no cromossomo I_b. O alelo *R* está localizado no cromossomo II_a e o alelo *r* está localizado na mesma posição no cromossomo II_b.
a. Desenhe estes cromossomos, identificando os genes *M*, *m*, *P*, *p*, *R* e *r*, à medida que eles podem aparecer na metáfase I da meiose. Considere que não há *crossing over*.
b. Levando em consideração a separação aleatória dos cromossomos na anáfase I, desenhe os cromossomos (com os genes identificados) encontrados em todos os possíveis tipos de gametas que possam surgir a partir desta célula que está passando pela meiose. Considere que não há *crossing over*.

35. Um cavalo tem 64 cromossomos e um jumento, 62. Um cruzamento entre uma égua e um jumento produz uma mula, que geralmente é estéril. Quantos cromossomos uma mula tem? Você consegue pensar nos motivos de a maioria das mulas ser estéril?

*36. As células somáticas normais dos cavalos têm 64 cromossomos ($2n = 64$). Quantos cromossomos e moléculas de DNA estarão presentes nos seguintes tipos de células de cavalo?

(Tamara Didenko/iStockphoto.)

Tipo de célula	Número de cromossomos	Número de moléculas de DNA
a. Espermatogônia	_____	_____
b. Primeiro corpo polar	_____	_____
c. Ovócito primário	_____	_____
d. Espermatócito secundário	_____	_____

37. Indique se cada uma das células a seguir é haploide ou diploide.

Tipo de célula	Haploide ou diploide?
Microesporo	_____
Espermatócito primário	_____
Microesporócito	_____
Primeiro corpo polar	_____
Ovogônia	_____
Espermátide	_____
Megaesporo	_____
Óvulo	_____
Ovócito secundário	_____
Espermatogônia	_____

*38. Um ovócito primário se divide para dar origem a um ovócito secundário e um primeiro corpo polar. O ovócito secundário, então, divide-se para originar um óvulo e um segundo corpo polar.
   a. As informações genéticas encontradas no primeiro corpo polar são idênticas às encontradas no ovócito secundário? Explique sua resposta.
   b. As informações genéticas encontradas no segundo corpo polar são idênticas às encontradas no óvulo? Explique sua resposta.

Questões desafiadoras

Seção 2.3

39. Oitenta a 90% das anomalias cromossômicas mais comuns nos seres humanos surgem porque os cromossomos não se dividem corretamente na ovogênese. Você consegue imaginar porque a falha na divisão dos cromossomos seria mais comum na gametogênese feminina do que na gametogênese masculina?
40. Na média, qual proporção do genoma nos pares de seres humanos a seguir seria exatamente a mesma se não ocorresse *crossing over*? (Para esta pergunta, vamos ignorar o caso especial dos cromossomos X e Y e considerar que todos os genes estão localizados nos cromossomos não sexuais).
   a. Pai e filho.
   b. Mãe e filho.
   c. Dois irmãos (descendentes que têm mesmo pai e mãe).
   d. Meio-irmãos (descendentes que têm apenas pai ou mãe em comum).
   e. Tio e sobrinho.
   f. Avô e neto.

*41. As abelhas-fêmeas são diploides e os machos são haploides. Os machos haploides produzem espermatozoides e cruzam de modo bem-sucedido com as fêmeas diploides. Os ovos fertilizados produzem fêmeas e os ovos não fertilizados produzem

machos. Você acha que o processo de produção de espermatozoides nas abelhas-macho é diferente da produção em outros animais?

*Esta analogia é adaptada de K. Nasmyth. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annual Review of Genetics* 35:673-745, 2001.

3 Princípios Básicos da Hereditariedade

A genética do cabelo ruivo

Por ser um tom exótico ou diferente, o cabelo ruivo sempre foi motivo de fascínio para historiadores, poetas, artistas e cientistas. Os historiadores fizeram uma observação especial sobre o fato de que Boudica, a rainha céltica que liderou uma revolta contra o Império Romano, tinha um "grande volume de cabelo ruivo". Os primeiros artistas cristãos frequentemente retratavam Maria Madalena com um impressionante cabelo ruivo (embora não exista menção na Bíblia) e o famoso artista Botticelli pintou a deusa Vênus como uma beleza de cabelo ruivo na sua obra-prima *O Nascimento de Vênus*. A Rainha Elizabeth I da Inglaterra tinha cabelo ruivo crespo e, durante seu reinado, esta cor era a moda na sociedade londrina.

A cor do nosso cabelo provém em grande parte por um pigmento chamado melanina, que surge em duas formas principais: eumelanina, que é o preto ou marrom e feomelanina, que é vermelho ou amarelo. A cor do cabelo de uma pessoa é determinada por dois fatores: (1) a quantidade de melanina produzida (quanto mais melanina mais escuro o cabelo, quanto menos melanina, mais claro o cabelo) e (2) as quantidades relativas de eumelanina e feomelanina (mais eumelanina produz cabelo preto ou marrom, mais feomelanina produz cabelo ruivo ou loiro). A cor do nosso cabelo não é apenas uma curiosidade de pesquisa; a melanina protege contra os efeitos prejudiciais da luz do sol, e as pessoas ruivas têm uma pele mais sensível, particularmente ao câncer de pele.

A hereditariedade do cabelo ruivo tem sido assunto de debate científico. Em 1909, Charles e Gertrude Davenport especularam sobre a hereditariedade da cor do cabelo nos seres humanos. Charles Davenport foi um dos primeiros entusiastas da genética, em particular sobre a herança nos seres humanos, e foi o primeiro diretor do Laboratório de Biologia em Cold Spring Harbor, Nova Iork. Posteriormente, ele se tornou um proponente da eugenia, um movimento – hoje em dia desacreditado – que defendia o aprimoramento da raça humana por intermédio da genética. O estudo dos Davenport tinha como base as histórias familiares enviadas por amadores sem treinamento, e sua metodologia era falha, mas seus resultados sugeriram que o cabelo ruivo era recessivo em relação ao preto e marrom, o que significa que uma pessoa precisava herdar duas cópias de um gene de cabelo ruivo – um do pai e um da mãe – para ter o cabelo ruivo. No entanto, pesquisas posteriores contradisseram essa conclusão inicial, sugerindo que cabelo ruivo era herdado, ao contrário, como uma característica dominante e que a pessoa teria o cabelo ruivo apenas se tivesse um único gene para cabelo ruivo. A controvérsia se a cor vermelha do cabelo é dominante ou recessiva ou dependente de combinações de vários genes diferentes continuou por muitos anos.

Em 1993, os cientistas que investigaram um gene que afeta a cor do pelo dos camundongos descobriram que o gene codifica o receptor 1 de melanocortina. Este receptor, quando ativado, aumenta a produção de eumelanina preta e reduz a produção da feomelanina vermelha, resultando no pelo preto ou marrom. Pouco depois, o mesmo gene do receptor 1 de melanocortina (*MC1R*) foi localizado no cromossomo 16 humano e foi analisado. O cabelo ruivo surge quando este gene sofre mutação nos seres humanos. A maior parte das pessoas com cabelo ruivo carrega duas cópias defeituosas do gene *MC1R*, o que significa que a característica é recessiva (como proposto originalmente pelos Davenport em 1909). Entretanto, de 10 a 20% das pessoas ruivas têm apenas uma única cópia mutante de *MC1R*, confundindo a interpretação recessiva do cabelo ruivo (as pessoas com uma única cópia mutante do gene tendem a ter cabelo ruivo mais claro do que aquelas com duas cópias mutantes). O tipo e a frequência de mutações no gene *MC1R* variam muito entre as populações de seres humanos, sendo responsáveis pelas diferenças étnicas na preponderância do cabelo ruivo: nas pessoas descendentes da África e Ásia, as mutações de cabelo ruivo são raras, enquanto quase 40 % das pessoas na parte norte do Reino Unido possuem pelo menos uma cópia mutante do gene para cabelo ruivo.

Os seres humanos modernos não são os únicos ruivos. A análise de DNA de ossos de ancestrais indica que alguns Neandertais também carregavam uma mutação no gene *MC1R* que, provavelmente, desenvolvia o cabelo ruivo, mas a mutação é diferente da observada nos seres humanos modernos.

ste capítulo estuda os princípios da hereditariedade: como os genes – como os indicados para o receptor 1 da melanocortina

E– são passados de geração para geração e como os fatores como dominância influenciam a herança. Os princípios da hereditariedade foram desenvolvidos primeiro por Gregor Mendel, então começamos este capítulo examinando suas descobertas científicas. Assim, iremos para os cruzamentos genéticos simples, nos quais uma única característica é examinada. Consideraremos algumas técnicas para prever o desfecho dos cruzamentos genéticos e então vamos examinar os cruzamentos em duas ou mais características. E veremos como os princípios aplicados a cruzamentos genéticos simples e as razões de descendentes que eles produzem servem como a chave para compreender os cruzamentos mais complicados. O capítulo termina como uma discussão dos testes estatísticos para analisar os cruzamentos.

Vários conceitos vão se misturar por todo o capítulo: os princípios da segregação e da segregação independente de Mendel, a probabilidade e o comportamento dos cromossomos. No primeiro momento parece que estes conceitos não estão relacionados, mas, na verdade, são diferentes visões do mesmo fenômeno porque os genes que sofrem segregação e segregação independente estão localizados nos cromossomos. Este capítulo examina as diferentes visões e esclarece suas relações.

3.1 Gregor Mendel descobriu os princípios básicos da hereditariedade

Em 1909, quando os Davenport especularam sobre a herança do cabelo ruivo, os princípios básicos da hereditariedade estavam apenas começando a serem descobertos pelos biólogos. Surpreendentemente, estes princípios foram descobertos 44 anos antes por Gregor Johann Mendel (1822-1884; **Figura 3.1**).

Mendel nasceu onde agora é a República Tcheca. Embora seus pais fossem simples fazendeiros com poucos recursos, ele recebeu uma educação sólida e foi aceito no monastério Augustianino em Brno em setembro de 1843. Após terminar o seminário, Mendel for ordenado padre e indicado para o cargo de professor em uma escola local. Ele se destacou no magistério e o abade do monastério o recomendou para estudar na Universidade de Viena, que ele frequentou de 1851 a 1853. Na Universidade, Mendel se inscreveu no recém-inaugurado Instituto de Física e assistiu aulas de matemática, química, entomologia, paleontologia, botânica e fisiologia vegetal. Foi provavelmente neste local onde Mendel conheceu o método científico, que mais tarde aplicou com tanto sucesso nos seus experimentos genéticos. Após 2 anos de estudos em Viena, Mendel retornou para Brno, onde lecionou e começou seu trabalho experimental com ervilhas. Ele realizou experimentos de procriação de 1956 a 1863 e apresentou seus resultados publicamente em reuniões na Brno Natural Science Society em 1865. O artigo que Mendel escreveu a partir destas palestras foi publicado em 1866. Entretanto, apesar do frequente interesse em hereditariedade, o efeito de sua pesquisa na comunidade científica foi mínimo. Naquela época, parece que ninguém observou que Mendel tinha descoberto os princípios básicos da hereditariedade.

Figura 3.1 Gregor Johann Mendel, ao fazer experimentos com ervilhas, foi o primeiro a descobrir os princípios da hereditariedade. (James King-Holmes/Photo Researchers.)

Em 1868, Mendel foi eleito abade do seu mosteiro e o aumento nas suas obrigações administrativas não permitiu que ele continuasse a lecionar e, consequentemente, a fazer seus experimentos genéticos. Ele morreu aos 61 anos em 6 de janeiro de 1884, sem o reconhecimento da sua contribuição para a genética.

A importância da descoberta de Mendel não foi reconhecida até 1900, quando três botânicos – Hugo de Vries, Erich von Tschermak-Seysenegg e Carl Correns – começaram a fazer independentemente experimentos semelhantes com plantas e chegaram a conclusões semelhantes às de Mendel. Ao descobrirem o artigo de Mendel, eles interpretaram seus resultados de acordo com seus princípios e atraíram a atenção para seu trabalho pioneiro.

O sucesso de Mendel

A abordagem de Mendel para o estudo da hereditariedade foi efetiva por vários motivos. Primeiro, foi sua escolha do material a ser estudado, a ervilha *Pisum sativum* (**Figura 3.2**), que oferecia claras vantagens para a investigação genética. Era uma planta de fácil cultivo e Mendel tinha o jardim e a estufa do monastério a sua disposição. Comparada com outras plantas, a ervilha cresce com relativa facilidade, completando uma geração inteira em uma única estação de crescimento. Para os padrões atuais, uma geração por ano parece ser muito lenta – as moscas-da-fruta completam uma geração em 2 semanas, e as bactérias em 20 min – mas Mendel não estava sob pressão para publicar rapidamente e foi capaz de seguir a herança das características individuais por várias gerações. Se ele tivesse escolhido trabalhar com um organismo com período de geração maior – cavalos, por exemplo – poderia nunca ter descoberto a base da herança. A ervilha também produz muitos descendentes – suas sementes – o que permitiu que Mendel detectasse razões matemáticas significativas nas características que observava nos descendentes.

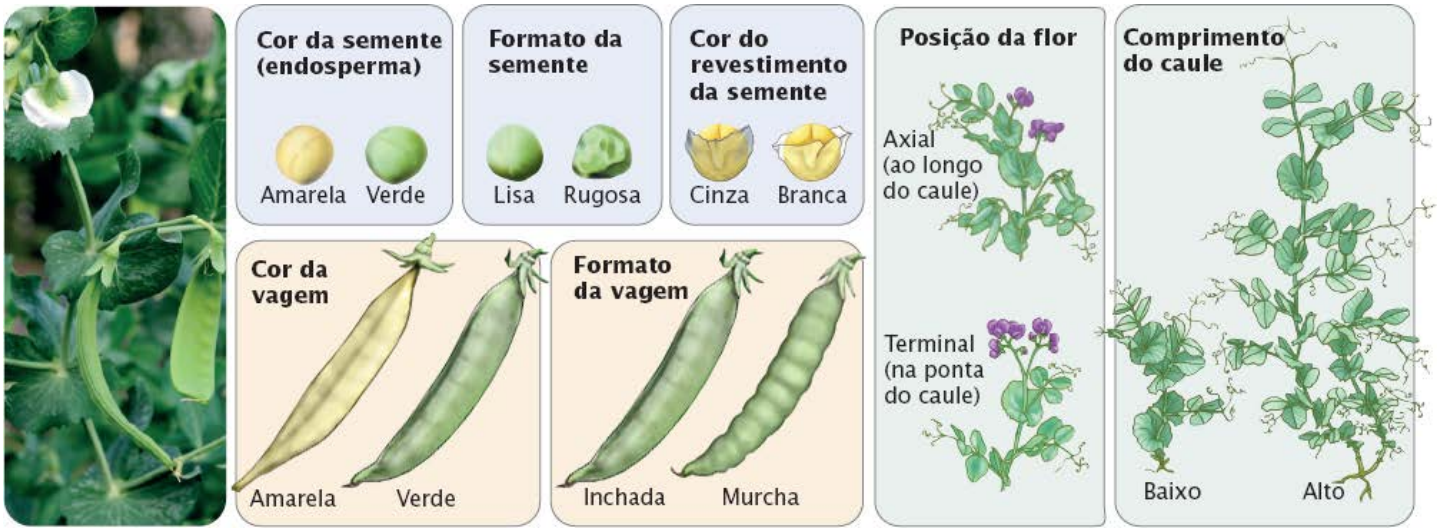


Figura 3.2 Mendel usou a ervilha *Pisum sativum* nos seus estudos de hereditariedade. Ele examinou 7 características que apareciam nas sementes e nas plantas que cresciam a partir das sementes. (Fotografia de Charles Stirling/Alamy.)

O grande número de variedades de ervilhas que Mendel tinha disponível foi crucial porque estas variedades tinham muitas características diferentes e eram geneticamente puras. Mendel, portanto, foi capaz de iniciar com plantas de composição genética variável e conhecida.

Boa parte do sucesso de Mendel pode ser atribuída às sete características que ele escolheu para estudar (ver Figura 3.2). Ele evitou características que tinham grande variação, e, pelo contrário, concentrou sua atenção nas que existiam em duas formas de fácil diferenciação, como o revestimento branco *versus* cinza da semente, sementes lisas *versus* rugosas e vagem inchada *versus* murcha.

Finalmente, Mendel teve sucesso porque ele adotou uma abordagem experimental e interpretou seus resultados usando matemática. Diferente de vários investigadores anteriores que apenas descreveram os *resultados* dos cruzamentos, Mendel formulou *hipóteses* com base nas suas observações iniciais e, então, realizou cruzamentos adicionais para testar suas hipóteses. Ele manteve cuidadosamente os registros de várias gerações, contendo cada tipo de traço, e calculou razões dos diferentes tipos. Ele era especialista em observar os padrões em detalhes, paciente e minuciosamente, conduzindo seus experimentos por 10 anos antes de tentar escrever seus resultados. **Resolva o Problema 13**

Conceitos

Gregor Mendel estabeleceu os princípios básicos da herança e publicou suas descobertas em 1866. Boa parte do sucesso de Mendel pode ser atribuída a 7 características que ele estudou.

Checagem dos conceitos 1

Qual dos fatores a seguir não contribuiu para o sucesso de Mendel no seu estudo da hereditariedade?

a. Ele usou ervilha.
b. Ele estudou os cromossomos de plantas.
c. Ele adotou uma abordagem experimental.
d. Ele usou matemática.

Terminologia genética

Antes de examinarmos os cruzamentos de Mendel e as conclusões que ele tirou destes experimentos, vamos revisar alguns termos muito usados em genética que serão úteis (**Quadro 3.1**). *Gene* é uma palavra que Mendel não conheceu. Ela não tinha sido criada até 1909, quando o geneticista dinamarquês Wilhelm Johannsen usou pela primeira vez. A definição de gene varia com o contexto de seu uso, e então sua definição varia à medida que exploramos diferentes aspectos da hereditariedade. Para nosso uso no contexto dos cruzamentos genéticos, definimos um **gene** como um fator herdado que determina uma característica.

Quadro 3.1 Resumo dos termos genéticos importantes.

Termo	Definição

Gene	Um fator herdado (região do DNA) que ajuda a determinar uma característica
Alelo	Uma ou mais formas alternativas de um gene
Locus	Local específico em um cromossomo ocupado por um alelo
Genótipo	Conjunto de alelos pertencentes a um organismo
Heterozigoto	Um organismo que tem dois alelos diferentes em um *locus*
Homozigoto	Um organismo que tem dois alelos iguais em um *locus*
Fenótipo ou traço	A aparência ou manifestação de uma característica
Característica ou caráter	Um atributo ou característica de um organismo

É comum os genes virem em diferentes versões chamadas **alelos** (**Figura 3.3**). Nos cruzamentos de Mendel, o formato da semente foi determinado por um gene que existe como dois diferentes alelos: um alelo codifica as sementes lisas e o outro codifica as sementes rugosas. Todos os alelos para determinado gene serão encontrados em um local específico no cromossomo chamado de ***locus*** para aquele gene (o plural de *locus* é *loci*). Assim, existe um local específico – um *locus* – em um cromossomo nas ervilhas onde o formato das sementes é determinado. Este *locus* pode ser ocupado por um alelo para sementes lisas ou para sementes rugosas. Usaremos o termo *alelo* para uma versão específica de um gene, e o termo *gene* para citar de forma mais geral qualquer alelo em um *locus*.

O **genótipo** é o conjunto de alelos de um organismo. Um organismo diploide com um genótipo que consiste em 2 alelos idênticos é **homozigoto** para aquele *locus*. Um organismo diploide que tem um genótipo com dois diferentes alelos é **heterozigoto** para aquele *locus*.

Outro termo importante é **fenótipo**, que é a manifestação ou aparência de uma característica. Um fenótipo pode se referir a qualquer tipo de característica – física, psicológica, bioquímica ou comportamental. Assim, a condição de sementes lisas é um fenótipo, um peso corporal de 50 quilos (50 kg) é um fenótipo e a anemia falciforme é um fenótipo. Neste livro, o termo *característica* ou *caráter* refere-se à característica geral como a cor dos olhos; o termo *traço* ou *fenótipo* se refere a manifestações específicas desta característica, como olhos castanhos ou azuis.

Um dado fenótipo surge a partir de um genótipo que se desenvolve em um ambiente específico. O genótipo determina o potencial para o desenvolvimento, ele define alguns limites, ou fronteiras, nesse desenvolvimento. Como o fenótipo se desenvolve dentro destes limites, ele é determinado pelos efeitos de outros genes e fatores ambientais, e o equilíbrio entre esses efeitos varia de característica a característica. Para algumas características, as diferenças entre os fenótipos são determinadas em grande parte pelas diferenças nos genótipos. Nas ervilhas de Mendel, por exemplo, o genótipo, e não o ambiente, determinava em grande parte o formato das sementes. Para outras características, as diferenças ambientais são mais importantes. A altura alcançada por um carvalho na sua maturidade é um fenótipo fortemente influenciado por fatores ambientais, como a disponibilidade de água, luz do sol e nutrientes. Contudo, o genótipo da árvore ainda impõe alguns limites na sua altura: um carvalho não cresce acima de 300 metros, não importa o quanto de luz do sol, água e fertilizante ele receba. Ou seja, até a altura de um carvalho é determinada em algum grau pelos genes. Para muitas características, os genes e o ambiente são importantes para determinar as diferenças fenotípicas.

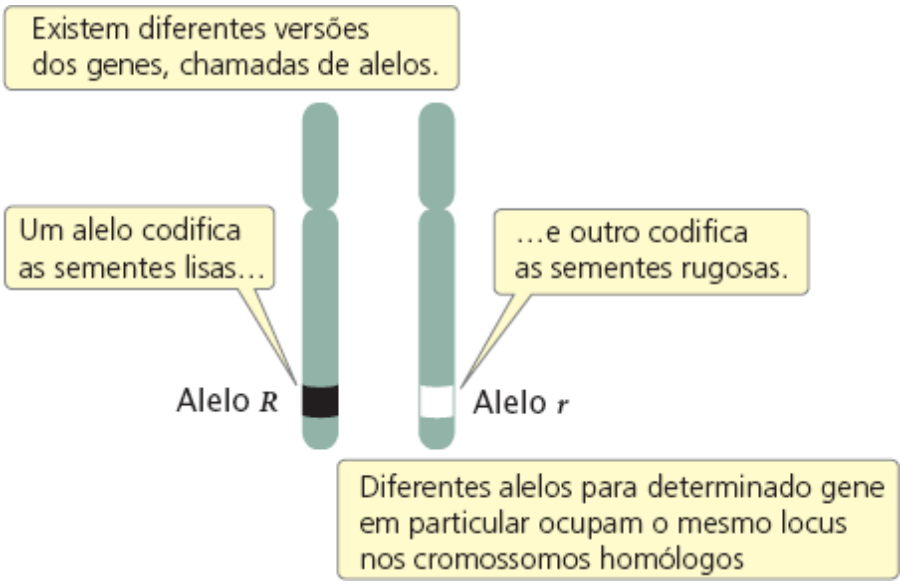


Figura 3.3 Em cada *locus*, um organismo diploide tem dois alelos localizados em diferentes cromossomos homólogos. Os alelos identificados aqui se referem aos traços estudados por Mendel.

Um conceito óbvio, mas importante, é que apenas os alelos do genótipo são herdados. Embora o fenótipo seja determinado, pelo menos em parte, pelo genótipo, os organismos não transmitem seus fenótipos para a próxima geração. A distinção entre genótipo e fenótipo é um dos princípios mais importantes da genética moderna. A próxima Seção descreve a observação cuidadosa do Mendel nos fenótipos por meio das várias gerações de experimentos de procriação. Estes experimentos permitiram que ele deduzisse não apenas os genótipos das plantas individuais, mas também as regras que governam sua hereditariedade.

Conceitos

Cada fenótipo resulta de um genótipo que se desenvolveu em um ambiente específico. Os alelos do genótipo e não o fenótipo são herdados.

Checagem dos conceitos 2

Qual é a diferença entre um *locus* e um alelo? Qual a diferença entre genótipo e fenótipo?

3.2 Os cruzamentos mono-híbridos revelam o princípio da segregação e o conceito de dominância

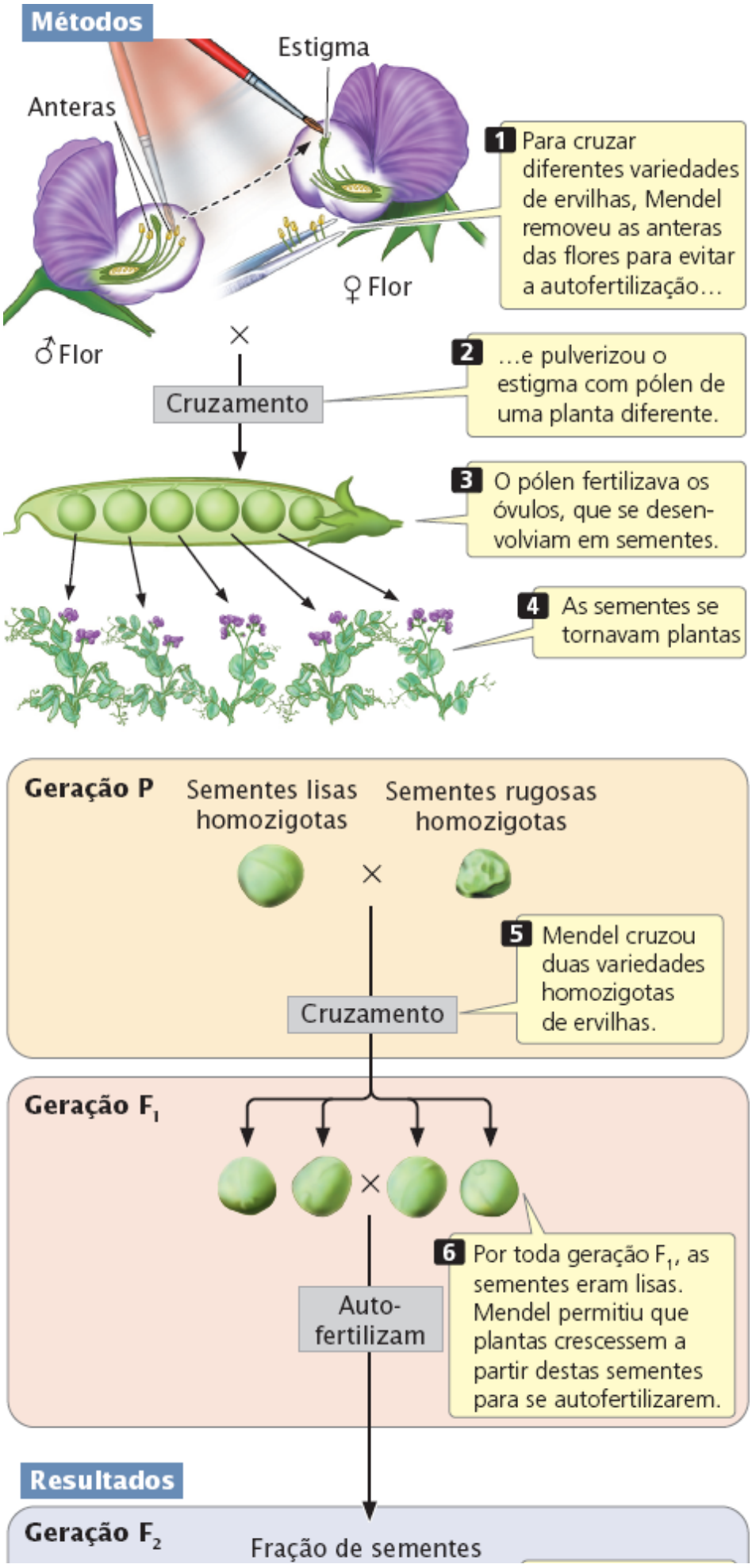
Mendel iniciou com 34 variedades de ervilhas e levou 2 anos selecionando as variedades que ele usaria nos experimentos. Ele verificou que cada variedade era geneticamente pura (homozigotos para cada um dos traços que ele escolheu estudar) ao criar as plantas por duas gerações e confirmar que os traços dos descendentes eram os mesmos que de seus genitores. Ele realizou vários cruzamentos entre as diferentes variedades. Embora as ervilhas façam a autofertilização (cada planta cruza com ela mesma), Mendel fez cruzamentos entre diferentes plantas ao abrir os botões antes que as anteras (órgãos sexuais masculinos) estivessem totalmente desenvolvidas, removia as anteras e então pulverizava o estigma (órgãos sexuais femininos) com pólen das anteras de uma planta diferente (**Figura 3.4**).

Mendel começou ao estudar os **cruzamentos mono-híbridos** – cujos genitores tinham apenas uma única característica diferente. Em um experimento, Mendel cruzou uma ervilha de geração pura (homozigota) para sementes lisas com uma que era pura para sementes rugosas (ver Figura 3.4). A primeira geração de um cruzamento é a **geração P (parental)**.

Após cruzar as duas variedades na geração P, Mendel observou a descendência do cruzamento. Com referência às características das sementes, como o formato da semente, o fenótipo se desenvolve assim que a semente amadurece porque os traços são determinados pelo embrião recém-formado na semente. Para as características associadas à própria planta, como comprimento do caule, o fenótipo não se desenvolve até a planta crescer a partir da semente; para estas características, Mendel tinha de esperar até a próxima geração, plantar as sementes e então observar os fenótipos nas plantas que germinaram.

Experimento

Pergunta: Quando as ervilhas com dois traços diferentes – sementes lisas e rugosas – são cruzadas, sua descendência exibe um desses traços, ambos os traços ou um traço intermediário?



Conclusão: Os traços das plantas genitoras não se combinam. Embora, as plantas de F_1 apresentem o fenótipo de um genitor, ambos os traços são transmitidos para a prole F_2 em uma proporção 3:1.

Figura 3.4 Mendel realizou cruzamentos mono-híbridos.

Os descendentes a partir dos genitores na geração P são a **geração F_1** (**primeira geração filial**). Quando Mendel examinou a geração F_1 deste cruzamento, descobriu que eles expressavam apenas um dos fenótipos presentes na geração original: todas as sementes F_1 eram lisas. Mendel realizou 60 desses cruzamentos e sempre tinha esse resultado. Ele também fez **cruzamentos recíprocos**: em um cruzamento, o pólen (gameta masculino) era retirado de uma planta com sementes lisas e em seu cruzamento recíproco, o pólen era retirado de uma planta com sementes rugosas. Os cruzamentos recíprocos também tinham o mesmo resultado: toda F_1 era lisa.

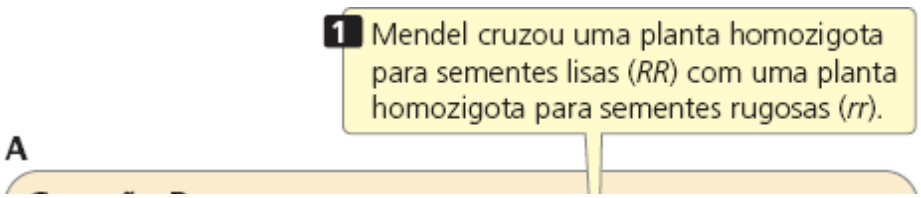
Mendel não ficou contente por examinar apenas as sementes que surgiam destes cruzamentos mono-híbridos. Na estação seguinte, ele plantou as sementes de F_1, cultivou as plantas que germinavam a partir destas sementes e permitiu que as plantas fizessem a autofertilização, produzindo **uma segunda geração filial** – a **geração F_2**. Ambos os traços da geração P surgiram na geração F_2, Mendel contou 5.474 sementes lisas e 1.850 sementes rugosas em F_2 (ver Figura 3.4). Ele observou que o número de sementes lisas e rugosas constituía uma proporção de aproximadamente 3:1, ou seja, 3/4 das sementes da geração F_2 era lisa e 1/4 era rugosa. Mendel conduziu cruzamentos monohíbridos para todas as sete características que ele estudou nas ervilhas e, em todos os cruzamentos ele obteve o mesmo resultado: todos lembravam um dos dois genitores, mas ambos os traços parentais em F_2 surgiam em uma razão aproximada de 3:1.

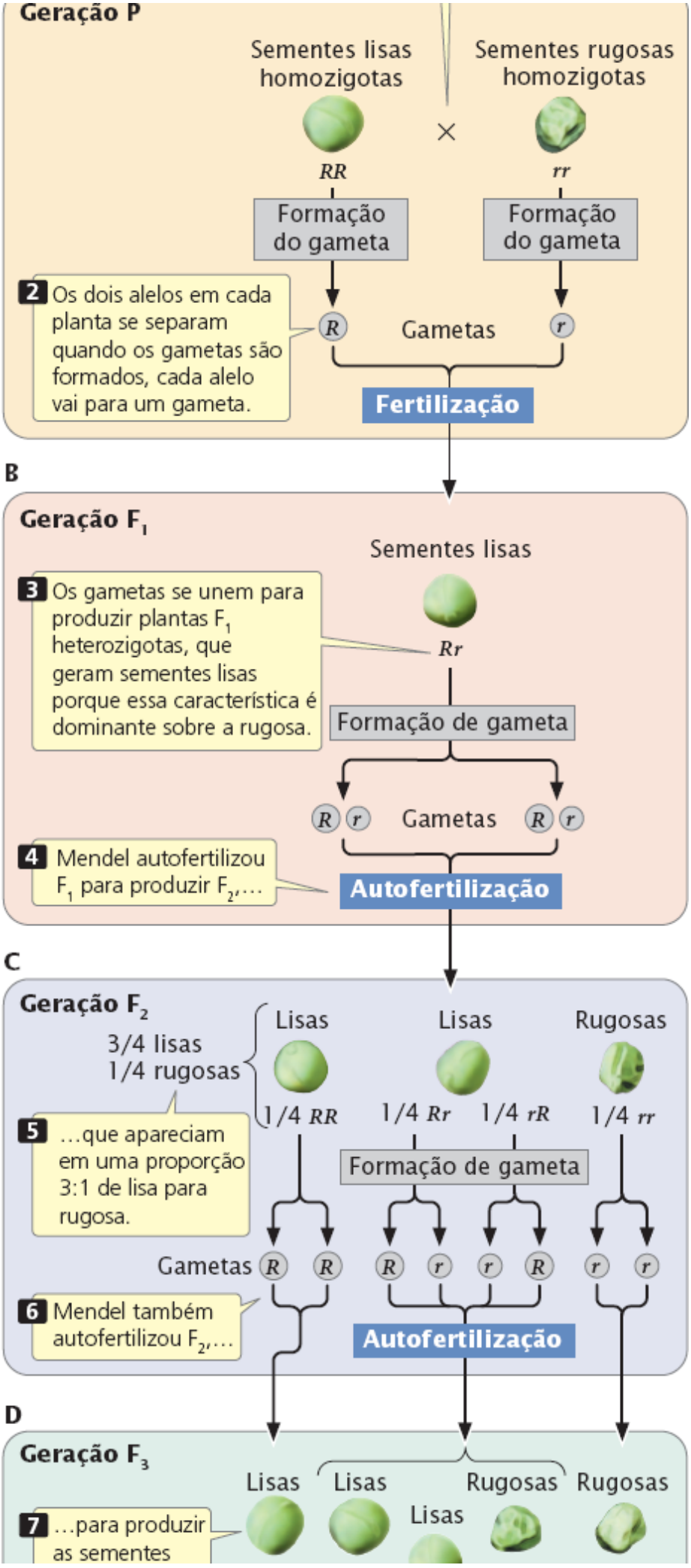
O que os cruzamentos mono-híbridos revelam

Primeiro, Mendel pensou que, embora as plantas de F_1 apresentassem o fenótipo de apenas um genitor, elas tinham obrigatoriamente herdado os fatores genéticos de ambos porque elas transmitiam ambos os fenótipos para a geração F_2. A existência de sementes lisas e rugosas nas plantas de F_2 poderia ser explicada apenas se as plantas F_1 tivessem os fatores genéticos para semente lisa e rugosa possivelmente herdados a partir da geração P. Ele concluiu que cada planta tinha obrigatoriamente dois fatores genéticos que codificam uma característica.

Os fatores genéticos (agora chamados alelos) que Mendel descobriu são, por convenção, designados por letras, o alelo para as sementes lisas é, geralmente, representado por um *R*, e o alelo para sementes rugosas, por *r*. As plantas na geração P dos cruzamentos de Mendel, tinham dois alelos idênticos: *RR* no genitor com semente lisa e *rr* no genitor com semente rugosa (**Figura 3.5 A**).

A segunda conclusão que Mendel tirou a partir dos cruzamentos mono-híbridos era que os dois alelos de cada planta se separavam quando os gametas eram formados, e cada alelo ia para um gameta. Quando dois gametas (um de cada genitor) se unem para produzir um zigoto, o alelo do genitor masculino se une com o alelo do genitor feminino para produzir o genótipo dos descendentes. Assim, as plantas F_1 de Mendel herdaram o alelo *R* da planta de semente lisa e um alelo *r* da planta de semente rugosa (**Figura 3.5 B**). Entretanto, apenas o traço codificado pelo alelo liso (R) *foi observado* em F_1: todos os descendentes de F_1 tinham sementes lisas. Mendel chamou esses traços que aparecem inalterados nos descendentes heterozigotos F_1 de **dominantes** e os traços que desapareceram nos descendentes heterozigotos ele chamou de **recessivos**. Os alelos para os traços dominantes são simbolizados com letras maiúsculas (p. ex., *R*), enquanto os alelos para os traços recessivos são simbolizados com letras minúsculas (p. ex., *r*). Quando os alelos dominantes e recessivos estão presentes juntos, o alelo recessivo é mascarado ou suprimido. O conceito de dominância foi a terceira conclusão importante que Mendel tirou a partir dos seus cruzamentos mono-híbridos.




Geração P
Sementes lisas homozigotas
Sementes rugosas homozigotas
RR
rr
Formação do gameta
Formação do gameta
2 Os dois alelos em cada planta se separam quando os gametas são formados, cada alelo vai para um gameta.
R
Gametas
r
Fertilização
B
Geração F1
Sementes lisas
3 Os gametas se unem para produzir plantas F1 heterozigotas, que geram sementes lisas porque essa característica é dominante sobre a rugosa.
Rr
Formação de gameta
R r
Gametas
R r
4 Mendel autofertilizou F1 para produzir F2,...
Autofertilização
C
Geração F2
3/4 lisas
1/4 rugosas
Lisas
Lisas
Rugosas
1/4 RR
1/4 Rr
1/4 rR
1/4 rr
5 ...que apareciam em uma proporção 3:1 de lisa para rugosa.
Formação de gameta
Gametas
R R
R r r R
r r
6 Mendel também autofertilizou F2,...
Autofertilização
D
Geração F3
Lisas
Lisas
Rugosas
Rugosas
Lisas
7 ...para produzir as sementes

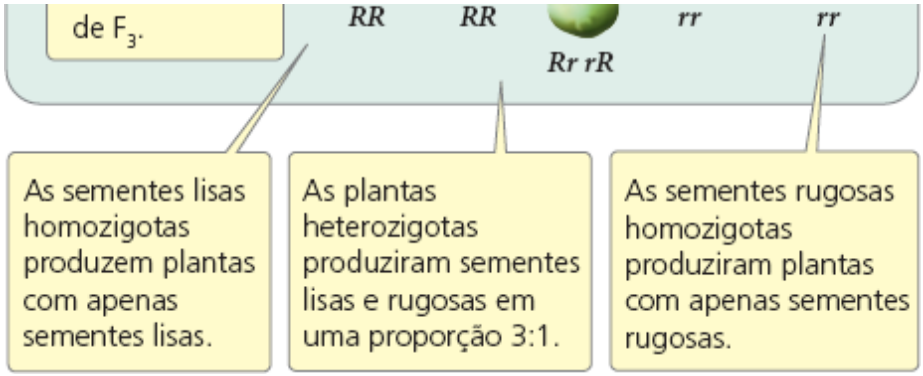


Figura 3.5 Os cruzamentos mono-híbridos de Mendel revelaram o Princípio da Segregação e o Conceito de Dominância.

A quarta conclusão de Mendel era de que dois alelos de uma planta se separam com a mesma probabilidade para os gametas. Quando as plantas de F_1 (com genótipo *Rr*) produziam os gametas, metade recebia o alelo *R* para as sementes lisas e metade recebia o alelo *r* para as sementes rugosas. Os gametas são aleatoriamente pareados para produzir os seguintes genótipos em proporções iguais em F_2: *RR, Rr, rR, rr* (**Figura 3.5 C**). Como liso (*R*) é dominante sobre rugoso (*r*), tivemos três descendentes lisos em F_2 (*RR, Rr, rR*) para cada descendente rugoso (*rr*) em F_2. Esta razão 3:1 dos descendentes de liso para rugoso que Mendel observou em F_2 poderia ser obtida apenas se os dois alelos de um genótipo se separassem nos gametas na mesma probabilidade.

As conclusões que Mendel descreveu sobre herança a partir dos seus cruzamentos mono-híbridos foram desenvolvidas posteriormente e formalizadas no princípio da segregação e no conceito de dominância. O **princípio da segregação** (primeira lei de Mendel, ver **Quadro 3.2**) afirma que cada organismo diploide individual tem dois alelos para uma característica em particular, um herdado do genitor materno e outro do genitor paterno. Esses dois alelos se segregam (separam) durante a formação dos gametas e cada um vai para um gameta. Além disso, os dois alelos se separam nos gametas em proporções iguais. O **conceito de dominância** afirma que quando existem dois alelos diferentes em um genótipo, apenas o traço codificado por um deles – o alelo "dominante" – é observado no fenótipo.

Mendel confirmou estes princípios ao permitir que suas plantas de F_2 se autofertilizassem e produzissem uma geração F_3. Ele descobriu que as plantas que cresciam a partir das sementes rugosas – as que apresentavam o traço recessivo (*rr*) produziam F_3 na qual todas as plantas produziam sementes rugosas. Como as plantas com sementes rugosas eram homozigotas para os alelos rugoso (*rr*), apenas estes alelos eram transmitidos para seus descendentes (**Figura 3.5 D**).

Quadro 3.2 Comparação dos princípios de segregação e segregação independente.

Princípio	Observação	Estágio da meiose*
Segregação (primeira lei de Mendel)	1. Cada organismo individual tem dois alelos que codificam um traço	Antes da meiose
	2. Os alelos se separam quando os gametas são formados	Anáfase I
	3. Os alelos se separam em proporções iguais	Anáfase I
Segregação independente (segunda lei de Mendel)	Os alelos em locais diferentes se separam de forma independente	Anáfase I

*Considerando que não ocorre *crossing over*. Se ocorrer *crossing over*, a segregação e a segregação independente também podem ocorrer na anáfase II da meiose.

As plantas crescem a partir das sementes lisas – o traço dominante – e caem em dois tipos (ver Figura 3.5 C). Na autofertilização, cerca de 2/3 destas plantas produzem sementes lisas e rugosas na geração F_3. Estas plantas eram heterozigotas (*Rr*); então elas produziram 1/4 *RR* (lisa), 1/2 *Rr* (lisa) e 1/4 *rr* (rugosa) de sementes, gerando uma proporção de 1:3 de sementes lisas e rugosas em F_3. Cerca de 1/3 das plantas cultivadas a partir das sementes lisas era do segundo tipo; elas produziram apenas o traço de semente lisa em F_3. Estas plantas eram homozigotas para o alelo liso (*RR*) e isto poderia produzir apenas descendentes de sementes lisas na geração F_3. Mendel plantou as sementes obtidas em F_3 e cuidou destas plantas por mais três

ciclos de autofertilização. Em cada geração, 2/3 das plantas de sementes lisas produzia uma geração lisa e rugosa, enquanto 1/3 produzia apenas a geração lisa. Estes resultados estavam totalmente consistentes com o princípio da segregação.

Conceitos

O princípio da segregação afirma que cada organismo tem dois alelos que podem codificar uma característica. Estes dois alelos se segregam durante a formação dos gametas e cada alelo vai para um gameta. O conceito de dominância afirma que quando dois alelos de um genótipo são diferentes, apenas o traço codificado por um deles – o alelo "dominante" – é observado no fenótipo.

Checagem dos conceitos 3

Como Mendel descobriu que cada ervilha carreava dois alelos que codificavam uma característica?

Conceitos conectantes

Como relacionar os cruzamentos genéticos com a meiose

Agora observamos como os resultados dos cruzamentos monoíbridos são explicados pelo princípio da segregação de Mendel. Muitos estudantes descobrem que eles gostaram de trabalhar com os cruzamentos genéticos, mas ficaram frustrados pela natureza abstrata dos símbolos. Talvez você se sinta assim também. E pode perguntar "O que estes símbolos representam? O que o genótipo *RR* significa com relação à biologia do organismo?". As respostas para estas perguntas dependem de relacionar os símbolos abstratos dos cruzamentos com a estrutura e o comportamento dos cromossomos, os depósitos das informações genéticas (ver *Capítulo 2*).

Em 1900, quando o trabalho de Mendel foi redescoberto e os biólogos começaram a aplicar seus princípios da hereditariedade, a relação entre genes e cromossomos ainda não estava clara. A teoria de que os genes estão localizados nos cromossomos (a **teoria cromossômica da herança**) foi desenvolvida no início do século 20 por Walter Sutton, um estudante da graduação na Universidade de Columbia. Por intermédio do estudo cuidadoso da meiose em insetos, Sutton documentou que cada par homólogo de cromossomos consiste em um cromossomo materno e um cromossomo paterno. Ao mostrar que estes pares se separam independentemente nos gametas durante a meiose, ele concluiu que este processo é a base biológica para os princípios de hereditariedade de Mendel. O citologista e embriologista alemão Theodor Boveri chegou a conclusões semelhantes quase simultaneamente.

Os símbolos usados nos cruzamentos genéticos, como *R* e *r*, são apenas notações para sequências específicas de DNA nos cromossomos que codificam fenótipos específicos. Os dois alelos do genótipo estão localizados em cromossomos diferentes, mas homólogos. Um cromossomo de cada par homólogo é herdado da mãe e o outro é herdado do pai. Na fase S da interfase meiótica, cada cromossomo se replica, produzindo duas cópias de cada alelo, um em cada cromátide (**Figura 3.6 A**). Os cromossomos homólogos se separam na anáfase I em dois alelos diferentes (**Figura 3.6 B e C**). Esta separação do cromossomo é a base do princípio da segregação. Na anáfase II da meiose, as duas cromátides de cada cromossomo replicado se separam, então cada gameta resultante da meiose carrega apenas um único alelo em cada *locus*, como o princípio da segregação de Mendel previa.

Se ocorrer *crossing over* na prófase I da meiose, então as duas cromátides de cada cromossomo replicado não são mais idênticas e a segregação de diferentes alelos ocorre na anáfase I e anáfase II (ver Figura 3.6 C). Entretanto, Mendel não sabia nada sobre os cromossomos; ele formulou seus princípios de hereditariedade totalmente com base nos resultados dos cruzamentos que fez. Contudo, não devemos esquecer que esses princípios funcionam porque se fundamentam no comportamento dos cromossomos reais na meiose. Resolva o Problema 30

A natureza molecular dos alelos

Vamos examinar com mais detalhes exatamente o que é um alelo e como ele determina um fenótipo. Embora Mendel não tivesse informações sobre a natureza física dos fatores genéticos nos seus cruzamentos, os geneticistas modernos determinaram agora a base molecular destes fatores e como eles codificam um traço como as ervilhas rugosas.

Os alelos, como *R* e *r* que codificam as ervilhas lisas e rugosas, geralmente representam sequências específicas de DNA. O *locus* que determina se uma ervilha é lisa ou rugosa é uma sequência de DNA no cromossomo 5 da ervilha que codifica uma

proteína chamada isoforma 1 da enzima ramificadora de amido (SBE-1). O alelo *R*, que produz sementes lisas nas ervilhas, codifica uma forma funcional, normal da enzima SBE-1. Esta enzima converte uma forma linear de amido em uma forma muito ramificada. O alelo *r*, que codifica as sementes rugosas, é uma sequência diferente de DNA que tem uma mutação ou erro, ela codifica uma forma não ativa da enzima que não produz a forma ramificada do amido e leva ao acúmulo de sacarose na ervilha *rr*. Como a ervilha *rr* tem muita sacarose, a semente em desenvolvimento absorve água e incha. Posteriormente, à medida que ela amadurece, perde água. Como as ervilhas *rr* absorvem mais água e se expandem mais durante o desenvolvimento, elas perdem mais água durante o amadurecimento e consequentemente parecem enrugadas ou rugosas. O alelo *r* para as sementes rugosas é recessivo porque a presença de um único alelo *R* no heterozigoto codifica enzima SBE-1 suficiente para produzir amido ramificado e sementes lisas.

A pesquisa revelou que o alelo *r* tem 800 pares de base extra do DNA que interrompem a sequência normal de codificação do gene. O DNA extra parece surgir de um elemento de transposição, um tipo de sequência de DNA que tem a capacidade de se deslocar de um local no genoma para outro, como vamos discutir adiante, no Capítulo 18.

Como prever os desfechos dos cruzamentos genéticos

Uma das metas de Mendel ao conduzir seus experimentos com ervilhas foi desenvolver uma forma para prever o desfecho dos cruzamentos entre as plantas com diferentes fenótipos. Nesta Seção, você aprenderá primeiro um método simples de notação para prever os desfechos dos cruzamentos genéticos (o quadrado de Punnett) e, então, você aprenderá como usar a probabilidade para prever os resultados dos cruzamentos.

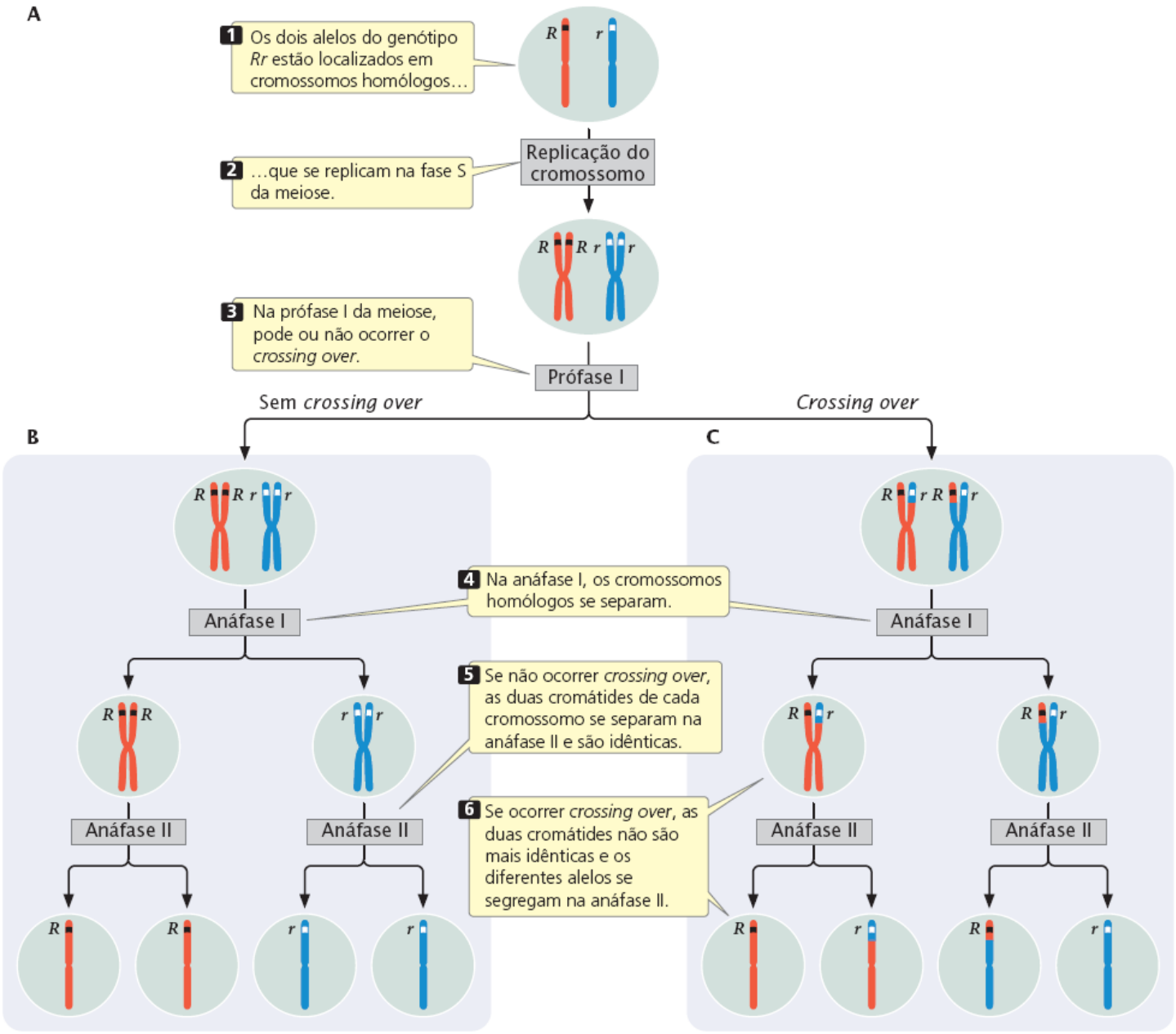


Figura 3.6 A segregação resulta da separação dos cromossomos homólogos na meiose.

O quadrado de Punnett. O quadrado de Punnett foi desenvolvido pelo geneticista inglês Reginald C. Punnett em 1917. Para ilustrar o quadrado de Punnett, vamos examinar outro cruzamento feito por Mendel. Ao cruzar duas variedades de ervilhas, com diferentes alturas, Mendel estabeleceu que a variedade alta (*T*) era dominante sobre a baixa (*t*). Ele testou sua teoria sobre a herança dos traços dominantes ao cruzar uma planta alta F_1 que era heterozigota (*Tt*) com a variedade baixa homozigota (*tt*). Este tipo de cruzamento, entre um genótipo F_1 e um dos genótipos genitores, é chamado de **retrocruzamento**.

Para prever os tipos de descendentes desse retrocruzamento, primeiro determinamos quais gametas serão produzidos por cada genitor (**Figura 3.7 A**). O princípio da segregação afirma que os alelos em cada genitor se separam e um alelo passa para cada gameta. Todos os gametas da planta baixa *tt* homozigota receberão um único alelo curto (*t*). A planta alta neste cruzamento é heterozigota (*Tt*): então 50% dos seus gametas receberão um alelo alto (*T*) e os outros 50% receberão um alelo baixo (*t*).

Um **quadrado de Punnett** é construído ao desenhar uma grade, colocando os gametas produzidos por um genitor ao longo da linha superior e os gametas produzidos pelo outro genitor no lado esquerdo (**Figura 3.7 B**). Cada célula (um bloco dentro do quadrado de Punnett) tem um alelo de cada gameta correspondente, gerando o genótipo dos descendentes produzidos pela fusão desses gametas. Na célula superior à esquerda do quadrado de Punnett na Figura 3.7 B, um gameta com *T* da planta alta se une com um gameta com *t* da planta baixa, gerando o genótipo do descendente (*Tt*). É útil escrever o fenótipo expresso por cada genótipo, aqui o descendente será alto, porque o alelo alto é dominante sobre o alelo baixo. Este processo é repetido para todas as células no quadrado de Punnett.

Ao contar, podemos determinar os tipos de descendentes produzidos e suas razões. Na Figura 3.7 B, duas células têm descendentes altos (*Tt*) e duas células tem descendentes baixos (*tt*), então a razão genotípica esperada para este cruzamento é 2 *Tt* para 2 *tt* (uma razão 1: 1). Outra forma de expressar este resultado é dizer que esperamos que 1/2 dos descendentes tenha genótipo *Tt* (e fenótipo de planta alta) e 1/2 dos descendentes tenha genótipo *tt* (e fenótipo de planta baixa). Neste cruzamento, a razão genotípica e a fenotípica são iguais, mas este desfecho pode não ser o caso. Tente preencher um quadrado de Punnett para o cruzamento no qual as plantas F_1 de semente lisa na Figura 3.5 sofrem autofertilização (você deve ter uma razão fenotípica de 3 sementes lisas para 1 rugosa e uma razão genotípica de 1 *RR* para 2 *Rr* para 1 *rr*).

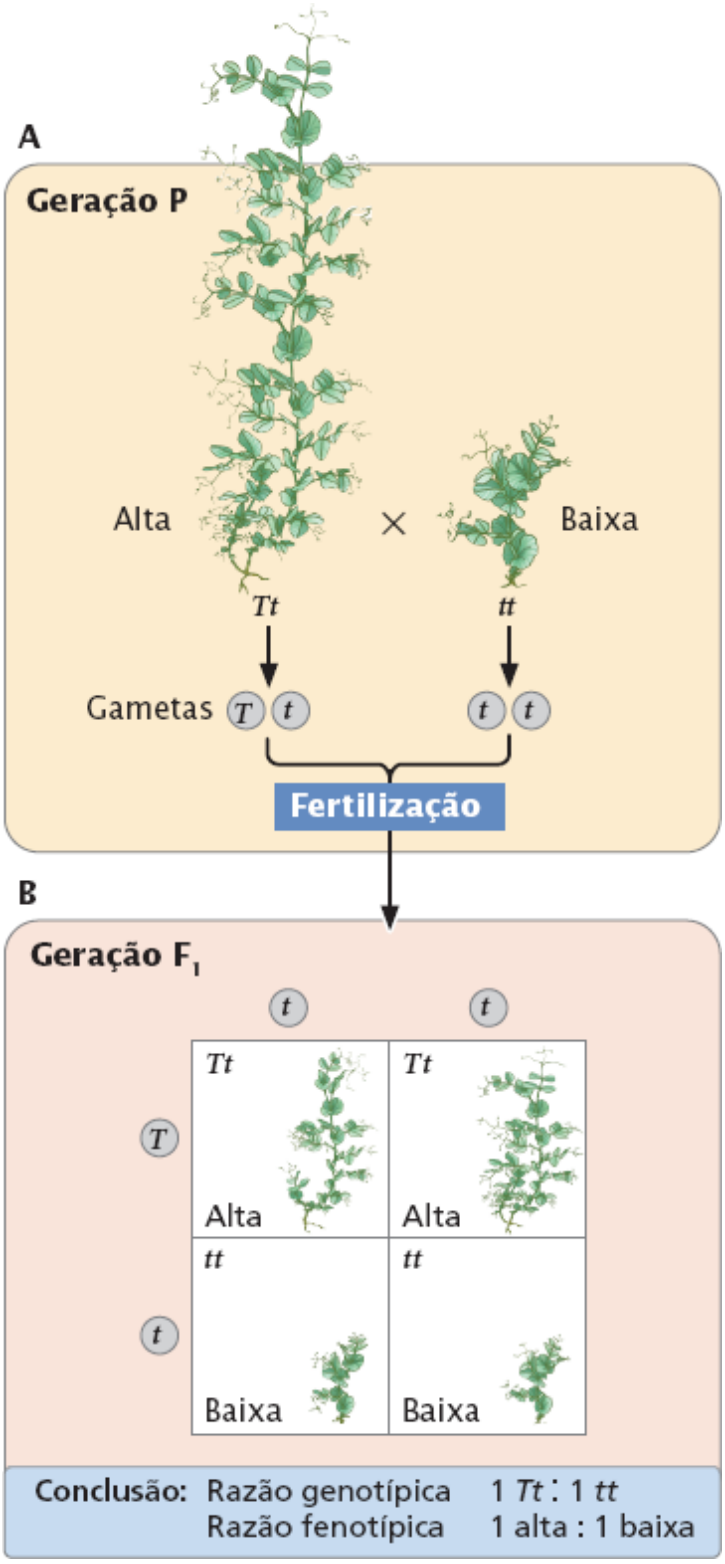


Figura 3.7 **O quadrado de Punnett pode ser usado para determinar os resultados de um cruzamento genético.**

Conceitos

O quadrado de Punnett é um método taquigráfico para prever as proporções de genótipo e fenótipo dos descendentes de um cruzamento genético.

Checagem dos conceitos 4

Se uma planta F_1 descrita na Figura 3.5 fizer um cruzamento reversível com um genitor com sementes lisas, qual proporção dos descendentes terá sementes rugosas? (Use um quadrado de Punnett.)

a. 3/4
b. 1/2
c. 1/4

d. 0

Probabilidade como ferramenta da genética. Outro método para determinar o desfecho de um cruzamento genético é usar as regras de probabilidade, como Mendel fez com seus cruzamentos. A **probabilidade** expressa quão provável é a ocorrência de determinado evento. É o número de vezes que um evento em particular ocorre, dividido pelo número de todos os possíveis desfechos. Por exemplo, um baralho de 52 cartas tem apenas um rei de copas. A probabilidade de retirar uma carta do baralho de forma aleatória e obter o rei de copas é de 1/52 porque existe apenas uma carta que é o rei de copas (um evento) e existem 52 cartas que podem ser retiradas do baralho (52 possíveis desfechos). A probabilidade de retirar uma carta e obter um ás é de 4/52 porque existem quatro cartas de ases (quatro eventos) e 52 cartas (possíveis desfechos). A probabilidade pode ser expressa como uma fração (4/52 neste caso) ou como um número decimal (0,077 neste caso).

A probabilidade de um evento específico é determinada ao saber algo sobre *como* e *com que frequência* o evento ocorre. Sabemos, por exemplo, que a probabilidade de jogar um dado de seis lados e obter um quatro é de 1/6 porque o dado tem seis lados e cada lado tem a mesma probabilidade de ficar no topo. Então, neste caso, a compreensão da natureza do evento, o formato do dado jogado, permite que determinemos a probabilidade. Em outros casos, determinamos a probabilidade de um evento ao fazer várias observações. Quando uma previsão do tempo diz que existe uma chance de 40% de chover em um dia específico, esta probabilidade foi obtida ao observar vários dias com condições atmosféricas semelhantes e descobrir que chove em 40% destes dias. Neste caso, a probabilidade foi determinada empiricamente (por observação).

A regra de multiplicação. Duas regras de probabilidade são úteis para prever as razões de descendentes produzidos nos cruzamentos genéticos. A primeira é a **regra de multiplicação**, que afirma que a probabilidade de dois ou mais eventos independentes ocorrerem juntos é calculada ao multiplicar suas probabilidades independentes.

Para ilustrar o uso da regra de multiplicação, vamos novamente jogar um dado. A probabilidade de jogar um dado e obter um quatro é 1/6. Para calcular a probabilidade de jogar duas vezes e obter 2 quatros, podemos aplicar a regra de multiplicação. A probabilidade de obter um quatro na primeira jogada é de 1/6 e a probabilidade de obter um quatro na segunda jogada é de 1/6, então a probabilidade de obter um quatro nas duas jogadas é 1/6 × 1/6 = 1/36 (**Figura 3.8 A**). O indicador-chave para aplicar a regra de multiplicação é a palavra *e* no exemplo que apresentamos; queremos saber a probabilidade de obter um quatro na primeira jogada *e* um quatro na segunda jogada.

Para que a regra de multiplicação seja válida, os eventos cuja probabilidade unida é calculada têm de ser independentes – o desfecho de um evento não deve influenciar o desfecho de outro. Por exemplo, o número que surge ao jogar o dado não tem influência no número que surge ao jogar outro dado; então estes eventos são independentes. Entretanto, se quisermos conhecer a probabilidade de levar uma pancada de martelo na cabeça e ir para o hospital no mesmo dia, não podemos apenas aplicar a regra de multiplicação e multiplicar os dois eventos juntos, porque eles não são independentes – levar uma pancada de martelo na cabeça certamente influencia a probabilidade de parar em um hospital.

A regra de adição. A segunda regra de probabilidade usada com frequência na genética é a **regra de adição**, que afirma que a probabilidade de um ou dois eventos mutuamente exclusivos é calculada ao adicionar as probabilidades destes eventos. Vamos examinar esta regra em questões práticas. Para obter a probabilidade de jogar um dado uma vez e obter *ou* um três *ou* um quatro, usaríamos a regra de adição, adicionando a probabilidade de obter um três (1/6) à probabilidade de obter um quatro (novamente 1/6) ou 1/6 + 1/6 = 2/6 = 1/3 (**Figura 3.8 B**). Os indicadores chaves para aplicar a regra de adição são as palavras *ou* e *ou*.

Para que a regra de adição seja válida, os eventos cuja probabilidade será calculada têm de ser mutuamente excludentes, o que significa que um evento exclui a possibilidade de ocorrência de outro evento. Por exemplo, você não pode jogar um único dado apenas uma vez e obter um três e um quatro, porque apenas um lado do dado estaria para cima. Estes eventos são mutuamente excludentes.

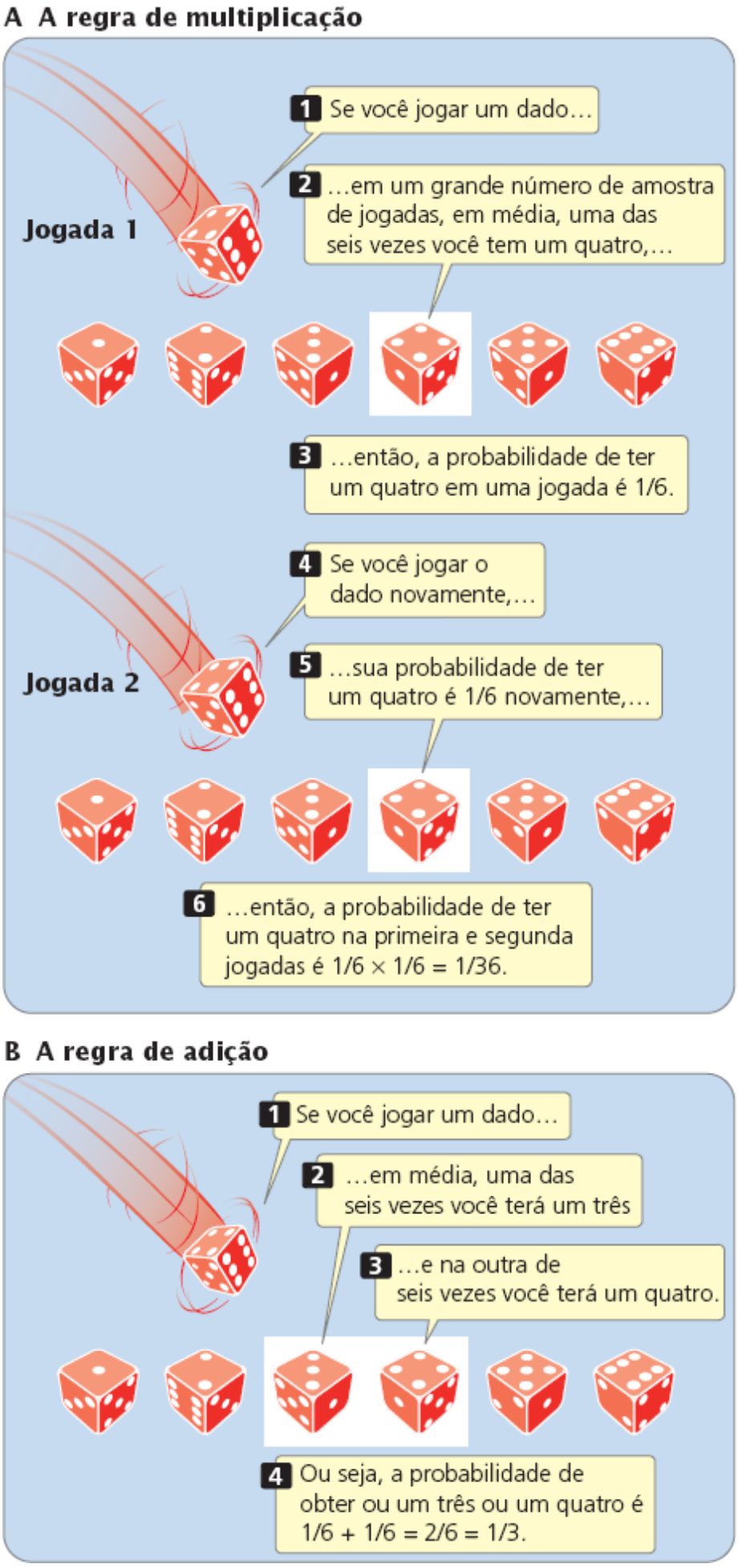


Figura 3.8 **As regras de multiplicação e adição podem ser usadas para determinar a probabilidade de combinações de eventos.**

Conceitos

A regra de multiplicação afirma que a probabilidade de dois ou mais eventos independentes ocorrerem juntos é calculada ao multiplicar suas probabilidades independentes. A regra de adição afirma que a probabilidade de um de dois ou mais eventos mutuamente excludentes ocorrer é calculada ao adicionar suas probabilidades.

Checagem dos conceitos 5

Se a probabilidade de ter o tipo sanguíneo A é 1/8 e a probabilidade de ter o tipo 0 é 1/2, qual é a probabilidade de ser tipo A ou tipo 0?

a. 5/8

b. 1/2

c. 1/10

d. 1/6

Aplicação da probabilidade nos cruzamentos genéticos. As regras de multiplicação e de adição da probabilidade podem ser usadas no lugar do quadrado de Punnett para prever as razões dos descendentes esperados a partir de um cruzamento genético. Primeiro, vamos considerar um cruzamento entre duas ervilhas heterozigotas para o *locus* que determina a altura, *Tt* × *Tt*. Metade dos gametas produzidos por cada planta terá um alelo *T* e a outra metade um alelo *t*, então a probabilidade para cada tipo de gameta é de 1/2.

Os gametas dos dois genitores podem se combinar de quatro formas para produzir os descendentes. Ao usar a regra de multiplicação, podemos determinar a probabilidade de cada possível tipo. Para calcular a probabilidade de obter uma geração *TT*, por exemplo, multiplicamos a probabilidade de receber um alelo *T* do primeiro genitor (1/2) vezes a probabilidade de receber um alelo *T* do segundo genitor (1/2). A regra de multiplicação deve ser usada aqui porque precisamos da probabilidade de receber um alelo *T* a partir do primeiro genitor e um alelo *T* do segundo genitor – dois eventos independentes. As quatro possibilidades de geração a partir deste cruzamento e suas probabilidades associadas são:

TT	(gameta *T* e gameta *T*)	1/2 × 1/2 = 1/2	alto
Tt	(gameta *T* e gameta *t*)	1/2 × 1/2 = 1/4	alto
tT	(gameta *t* e gameta *T*)	1/2 × 1/2 = 1/4	alto
tt	(gameta *t* e gameta *t*)	1/2 × 1/2 = 1/4	baixo

Observe que existem duas formas para produzir uma geração heterozigota: um heterozigoto pode receber um alelo *T* do primeiro genitor e um alelo *t* do segundo ou receber um alelo *t* do primeiro genitor e um alelo *T* do segundo.

Após determinar as probabilidades de obter cada tipo de geração, podemos usar a regra de adição para determinar as proporções gerais de fenótipos. Por causa da dominância, uma planta alta pode ter o genótipo *TT*, *Tt* ou *tT*; então, usando a regra de adição, encontramos a probabilidade da geração alta como sendo 1/4 + 1/4 + 1/4 = 3/4. Como apenas um genótipo codifica a planta baixa (*tt*), a probabilidade de descendentes baixos é de apenas 1/4.

Dois métodos foram introduzidos para solucionar os cruzamentos genéticos: o quadrado de Punnett e o método de probabilidade. Neste ponto você pode estar se perguntando “Por que se importar com regras de probabilidade e cálculos? O quadrado de Punnett é mais fácil e rápido também”. Isso é verdade para os cruzamentos mono-híbridos simples. Entretanto, para lidar com cruzamentos mais complexos que envolvem genes em dois ou mais *loci*, o método de probabilidade é mais seguro e rápido que o quadrado de Punnett.

A probabilidade condicional. Até aqui usamos a probabilidade para prever as chances de produzir certos tipos de descendentes considerando apenas os genótipos dos genitores. Às vezes, temos informações adicionais que modificam ou “condicionam” a probabilidade, uma situação chamada de **probabilidade condicional**. Por exemplo, imagine que cruzamos duas ervilhas heterozigotas (*Tt* × *Tt*) e obtemos uma planta alta. Qual é a probabilidade de que essa planta seja heterozigota (*Tt*)? Você pode imaginar que a probabilidade seria 1/2, a mesma de obter um descendente heterozigoto em um cruzamento entre dois heterozigotos. Entretanto, neste caso, temos algumas informações adicionais – o fenótipo da planta descendente – que modifica a probabilidade. Quando dois indivíduos heterozigotos são cruzados, esperamos uma geração de 1/4 *TT*, 1/2 *Tt* e 1/4 *tt*. Sabemos que os descendentes em questão são altos, então podemos eliminar a possibilidade de que tenha um genótipo *tt*. Os descendentes altos têm um genótipo *TT* ou genótipo *Tt* e em um cruzamento entre dois heterozigotos, eles ocorrem em uma razão 1:2. Portanto, a probabilidade de que uma geração alta seja heterozigota (*Tt*) é de dois em três, ou seja, 2/3.

A expansão binomial e a probabilidade. Quando a probabilidade é usada, é importante identificar que podem existir várias formas diferentes na qual um conjunto de eventos pode ocorrer. Considere dois genitores que sejam heterozigotos para o albinismo, uma condição recessiva nos seres humanos que provoca pigmentação reduzida na pele, cabelo e olhos (ver a introdução ao Capítulo 1). Quando dois pais heterozigotos para albinismo acasalam (*Aa* × *Aa*), a probabilidade de ter uma criança com albinismo (*aa*) é 1/4 e a probabilidade de ter uma criança com pigmentação normal (*AA* ou *Aa*) é 3/4. Suponha que queremos saber a probabilidade deste casal ter três crianças com albinismo. Neste caso, existe apenas uma forma de isto acontecer: seu primeiro filho tem albinismo *e* seu segundo filho tem albinismo *e* seu terceiro filho tem albinismo. Aqui, simplesmente aplicamos a regra de multiplicação: 1/4 × 1/4 × 1/4 = 1/64.

Imagine agora que queremos saber a probabilidade deste casal ter três crianças, uma com albinismo e duas com pigmentação normal. Esta situação é mais complicada. A primeira criança pode ter albinismo, enquanto a segunda e a terceira não são afetadas e a probabilidade desta sequência de eventos é 1/4 × 3/4 × 3/4 = 9/64. Por outro lado, a primeira e a terceira criança podem ter pigmentação normal, enquanto a segunda tem albinismo, a probabilidade desta sequência de eventos é 3/4 × 1/4 × 3/4 = 9/64. Finalmente as duas primeiras crianças podem ter pigmentação normal e a terceira ter albinismo, a probabilidade desta sequência de eventos é 3/4 × 3/4 × 1/4 = 9/64. Como *ou* a primeira sequência *ou* a segunda sequência *ou* a terceira sequência produzem uma criança com albinismo e duas com pigmentação normal, aplicamos a regra de adição e somamos as probabilidades: 9/64 + 9/64 +9/64 = 27/64.

Se quisermos saber a probabilidade de esse casal ter cinco crianças, duas com albinismo e três com pigmentação normal, imagine que *todas* as três combinações diferentes das crianças e suas probabilidades são mais difíceis. Esta tarefa é mais fácil se aplicarmos a expansão binomial.

A expansão binomial assume a forma $(p + q)^n$, em que p é igual à probabilidade de um evento, q é igual à probabilidade do evento alternativo e n é igual ao número de vezes que o evento ocorre. Para imaginar a probabilidade de duas das cinco crianças terem albinismo:

p = a probabilidade de uma criança ter albinismo (1/4)
q = a probabilidade de ter uma criança com pigmentação normal (3/4)

O binômio para esta situação é $(p + q)^5$ porque existem cinco crianças na família ($n = 5$). A expansão é:

$$(p + q)^5 = p^5 + 5p^4q + 10p^3q^2 + 10p^2q^3 + 5pq^4 + q^5$$

Cada um dos termos na expansão fornece a probabilidade para uma combinação em particular dos traços nas crianças. O primeiro termo na expansão (p^5) é igual à probabilidade de ter cinco crianças, todas com albinismo, porque p é a probabilidade de albinismo. O segundo termo ($5p^4q$) é igual à probabilidade de ter quatro crianças com albinismo e uma com pigmentação normal, o terceiro termo ($10p^3q^2$) é igual à probabilidade de ter três crianças com albinismo e duas com pigmentação normal e assim em diante.

Para obter a probabilidade de qualquer combinação dos eventos, inserimos os valores de p e q; então a probabilidade de ter duas de cinco crianças com albinismo é:

$$10p^2q^3 = 10(1/4)^2\ (3/4)^3 = 270/1.024 = 0{,}26.$$

Podemos descobrir com facilidade a probabilidade de qualquer combinação desejada do albinismo e pigmentação entre cinco crianças ao usar os outros termos na expansão.

Como expandimos este binômio neste exemplo? Em geral, a expansão de qualquer binômio $(p + q)^n$ consiste em uma série de termos $n + 1$. No exemplo anterior, $n = 5$, existem 5 + 1 = 6 termos; p^5, $5p^4q$, $10p^3q^2$, $10p^2q^3$, $5pq^4$ e q^5. Para escrever estes termos, primeiro imagine seus expoentes. O expoente de p no primeiro termo sempre começa com a potência para a qual o binômio é elevado, ou n. No nosso exemplo, n é igual a 5, então nosso primeiro termo é p^5. O expoente de p reduz por um em cada termo sucessivo, então o expoente de p é 4 no segundo termo (p^4), 3 no terceiro termo (p^3) e assim em diante. O expoente de q é 0 (sem q) no primeiro termo e aumenta em 1 em cada termo sucessivo, aumentando de 0 para 5 no nosso exemplo.

A seguir, determine o coeficiente de cada termo. O coeficiente do primeiro termo é sempre 1, então no nosso exemplo, o primeiro termo é $1p^5$ ou apenas p^5. O coeficiente do segundo termo é sempre o mesmo que a potência para o qual o binômio é elevado, em nosso exemplo, este coeficiente é 5 e o termo é $5p^4q$. Para o coeficiente do terceiro termo, olhe para o termo anterior, multiplique o coeficiente do termo anterior (5 no nosso exemplo) pelo expoente de p naquele termo (4) e então divida pelo número do termo (segundo termo, ou 2). Então, o coeficiente do terceiro termo em nosso exemplo é (5 × 4)/2 = 20/2 = 10 e o termo é $10p^3q^2$. Siga este procedimento para cada termo sucessivo. Os coeficientes para os termos na expansão binomial também podem ser determinados pelo triângulo de Pascal (**Quadro 3.3**). Os expoentes e coeficientes para cada termo nas cinco primeiras expansões binomiais são fornecidos no **Quadro 3.4**.

Outra forma para determinar a probabilidade de qualquer combinação específica de eventos é usar a seguinte fórmula:

$$P = \frac{n!\, p^s q^t}{s!t!}$$

em que P é igual a probabilidade geral do evento X, com a probabilidade *p* ocorrer *s* vezes, e o evento Y, com probabilidade de *q* ocorrer *t* vezes. Para nosso exemplo de albinismo, o evento X seria a ocorrência de uma criança com albinismo (1/4) e o evento Y seria a ocorrência de uma criança com pigmentação normal (3/4); *s* seria igual ao número de crianças com albinismo (2) e *t* seria igual ao número de crianças com pigmentação normal (3). O símbolo ! é o fatorial e significa o produto de todos os números inteiros de *n* a 1. Neste exemplo *n* = 5, então *n*! = 5 × 4 × 3 × 2 × 1. Ao aplicar esta fórmula para obter a probabilidade duas das cinco crianças terem albinismo, obtemos:

$$P = \frac{5!}{2!3!}(1/4)^2(3/4)^3$$
$$= \frac{5\times4\times3\times2\times1}{2\times1\times3\times2\times1}(1/4)^2(3/4)^3 = 0{,}26$$

Este valor é o mesmo obtido com a expansão binomial. **Resolva os Problemas 25 , 26 e 27**

Quadro 3.3 Triângulo de Pascal.

Os números de cada coluna representam os coeficientes de cada termo na expansão binomial $(p + q)^n$.

n	Coeficientes
	1
1	1 1
2	1 2 1
3	1 3 3 1
4	1 4 6 4 1
5	1 5 10 10 5 1
6	1 6 15 20 15 6 1

Nota: Cada número no triângulo, exceto para o primeiro, é igual à soma de dois números diretamente acima dele.

Quadro 3.4 Coeficientes e termos para a expansão binomial $(p + q)^n$ para $n = 1$ a 5.

n	Expansão binomial
1	$a + b$
2	$a^2 + 2ab + b^2$
3	$a^3 + 3a^2b + 3ab^2 + b^3$
4	$a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$
5	$a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$

O cruzamento-teste

Uma ferramenta útil para analisar os cruzamentos genéticos é o **cruzamento-teste**, no qual um indivíduo de genótipo desconhecido é cruzado com outro indivíduo com um genótipo homozigoto recessivo para o traço em questão. A Figura 3.7 ilustra um cruzamento-teste (neste caso, também é um retrocruzamento), que verifica ou revela o genótipo do primeiro indivíduo.

Imagine que você tenha uma ervilha de planta alta sem informações sobre seus genitores. Como a altura é um traço dominante nas ervilhas, sua planta pode ser homozigota (*TT*) ou heterozigota (*Tt*), mas você não saberia qual é. Você poderia determinar seu genótipo ao realizar um cruzamento-teste. Se a planta for homozigota (*TT*), um cruzamento-teste produziria todos os descendentes altos (*TT* × *tt* → todos *Tt*), se a planta for heterozigota (*Tt*), metade dos descendentes seria alta e metade seria baixa (*Tt* × *tt* → ½ *Tt* e ½ *tt*). Quando um cruzamento-teste é realizado, qualquer alelo recessivo no genótipo desconhecido é expresso nos descendentes, porque ele seria pareado com um alelo recessivo a partir do genitor homozigoto recessivo.

Resolva os Problemas 18 e 21

Conceitos

A expansão binomial pode ser usada para determinar a probabilidade de um conjunto particular de eventos. Um cruzamentoteste é um cruzamento entre um indivíduo com um genótipo desconhecido e um indivíduo com um genótipo homozigoto recessivo. O desfecho do cruzamento-teste pode revelar o genótipo desconhecido.

Símbolos genéticos

Como já observamos, os cruzamentos genéticos são, em geral, descritos com o uso de símbolos para indicar os diferentes alelos. Os símbolos usados para os alelos são geralmente determinados por uma comunidade de geneticistas que trabalham em um organismo particular, não havendo, portanto, sistema universal para indicá-los. Nas plantas, as letras minúsculas são usadas para indicar alelos recessivos e as letras maiúsculas são para os alelos dominantes. Duas ou três letras podem ser usadas para um único alelo: o alelo recessivo para as folhas em formato de coração nos pepinos é indicado por *hl* e o alelo recessivo para o formato anormal de cabeça de espermatozoide é *azh*.

Nos animais, o alelo comum para uma característica – chamada de **tipo selvagem** porque é o alelo geralmente encontrado no selvagem – é simbolizado por uma ou mais letras e um sinal (+). As letras são escolhidas com base no fenótipo mutante (incomum). Por exemplo, o alelo recessivo para os olhos amarelos na mosca da fruta oriental é representado por *ye*, enquanto o alelo para a cor dos olhos selvagem é representado por ye^+. Em alguns casos, as letras para o alelo selvagem são abandonadas e o alelo é representado apenas por um sinal de mais. Sobrescrito e subscrito também podem ser adicionados para diferenciar os genes: Lfr_1 e Lfr_2 representam os alelos dominantes mutantes em diferentes *loci* que produzem as margens da folha lacerada nas papoulas do ópio; El^R representa um alelo nas cabras que restringe o comprimento das orelhas.

Uma barra pode ser usada para distinguir os alelos presentes em um genótipo individual. Por exemplo, o genótipo de uma cabra que é heterozigota para orelhas restritas pode ser escrito como El^+/El^R ou simplesmente $+/El^R$. Se os genótipos em mais de um *locus* são apresentados juntos, um espaço separa os genótipos. Por exemplo, uma cabra heterozigota para um par de alelos que produz orelhas restritas e heterozigotas para outro par de alelos que produz bócio podem ser indicados por El^+/El^R *G/g*. Em alguns casos, é útil indicar a possibilidade de vários genótipos. Uma linha em um genótipo, como *A_*, indica que qualquer alelo é possível. Neste caso, *A_* pode incluir os genótipos *AA* e *Aa*.

Conceitos conectantes

Razões nos cruzamentos simples

Agora que já temos alguma experiência nos cruzamentos genéticos, vamos revisar as razões que aparecem na descendência de cruzamentos simples, nas quais um único *locus* está em análise e um dos alelos é dominante sobre outro. Compreender estas razões e os genótipos dos genitores que os produzem permitirá que você trabalhe os cruzamentos genéticos de forma rápida, sem recorrer ao quadrado de Punnett. Posteriormente, usaremos estas razões para trabalhar com cruzamentos mais complicados, que incluem vários *loci*.

Existem apenas três razões fenotípicas para compreender (**Quadro 3.5**). A razão 3:1 surge em um cruzamento genético simples quando ambos os genitores são heterozigotos para um traço dominante (*Aa* × *Aa*). A segunda razão fenotípica é 1:1, que resulta do acasalamento de um genitor heterozigoto com um genitor homozigoto. O genitor homozigoto neste cruzamento deve carrear dois alelos recessivos (*Aa* × *aa*) para obter uma razão 1:1, porque um cruzamento entre um genitor dominante homozigoto e um genitor heterozigoto (*AA* × *Aa*) produz uma geração que apresenta apenas o traço dominante.

A terceira razão fenotípica não é de fato uma razão: todos os descendentes têm o mesmo fenótipo (geração uniforme). Várias combinações de genitores podem produzir este desfecho (ver Quadro 3.5). Um cruzamento entre qualquer um dos dois genitores homozigotos – ou entre dois do mesmo homozigoto (*AA* ×

AA ou *aa* × *aa*) ou entre diferentes homozigotos (*AA* × *aa*) – produz uma geração tendo o mesmo fenótipo. A geração de um único fenótipo também pode surgir a partir de um cruzamento entre um genitor dominante homozigoto e um heterozigoto (*AA* × *Aa*).

Se estivermos interessados nas razões dos genótipos em vez dos fenótipos, existem apenas três desfechos para serem lembrados (**Quadro 3.6**): a razão 1:2:1, produzida por um cruzamento entre dois heterozigotos, a razão 1:1, produzida por um cruzamento entre um heterozigoto e um homozigoto, e descendentes uniformes, produzida por um cruzamento entre os dois homozigotos. Estas simples proporções de fenótipo e genótipo e os genótipos dos genitores que os produzem fornecem uma chave para compreender os cruzamentos para um único *locus* e, como você verá na Seção a seguir, para múltiplos *loci*.

Quadro 3.5 Razões fenotípicas para cruzamentos genéticos simples (cruzamentos para um único *locus*) com dominância.

Razão fenotípica	Genótipos dos genitores	Genótipos dos descendentes
3 : 1	*Aa* × *Aa*	3/4 *A_* : 1/4 *aa*
1 : 1	*Aa* × *aa*	1/2 *Aa* : 1/2 *aa*
Prole uniforme	*AA* × *AA*	Todos *AA*
	aa × *aa*	Todos *aa*
	AA × *aa*	Todos *Aa*
	AA × *Aa*	Todos *A_*

Quadro 3.6 Razões genotípicas para cruzamentos genéticos simples (cruzamentos para um único *locus*).

Razão genotípica	Genótipos dos genitores	Genótipos dos descendentes
1: 2: 1	*Aa* × *Aa*	1/4 *AA*: 1/2 *Aa*: 1/4 *aa*
1 : 1	*Aa* × *aa*	1/2 *Aa*: 1/2 *aa*
	Aa × *AA*	1/2 *Aa*: 1/2 *AA*
Prole uniforme	*AA* × *AA*	Todos *AA*
	aa × *aa*	Todos *aa*
	AA × *aa*	Todos *Aa*

3.3 Os cruzamentos di-híbridos revelam o princípio da segregação independente

Agora vamos estender o princípio da segregação de Mendel para cruzamentos mais complexos, que incluem alelos em múltiplos *loci*. Compreender a natureza destes cruzamentos requer um princípio adicional, o princípio da segregação independente.

Cruzamentos di-híbridos

Além do seu trabalho com cruzamentos mono-híbridos, Mendel cruzou variedades de ervilhas que tinham duas características diferentes – um **cruzamento di-híbrido**. Por exemplo, ele tinha uma variedade homozigota de ervilha com sementes lisas e amarelas, outra variedade homozigota com sementes rugosas e verdes. Quando ele cruzou as duas variedades, as sementes de todos os descendentes eram lisas e amarelas. Ele então fez a autofertilização e obteve a seguinte geração em F_2: 315 sementes lisas e amarelas; 101 sementes rugosas e amarelas; 108 sementes lisas e verdes e 32 sementes rugosas e verdes. Mendel reconheceu que estes traços apareciam em uma razão aproximadamente 9:3:3:1, ou seja, 9/16 dos descendentes eram lisas e amarelas; 3/16 eram rugosas e amarelas, 3/16 eram rugosas e verdes e 1/16 era rugosa e verde.

Princípio da segregação independente

Mendel fez vários cruzamentos di-híbridos para pares de características e sempre obtinha uma proporção 9:3:3:1 em F_2. Esta proporção faz sentido em relação à segregação e dominância se adicionarmos um terceiro princípio que Mendel identificou nos seus cruzamentos di-híbridos: o **princípio da segregação independente** (segunda lei de Mendel). Este princípio afirma que os alelos em diferentes *loci* se separam independentemente um do outro (ver Quadro 3.2).

Um erro comum é pensar que o princípio da segregação e o princípio da segregação independente se referem a dois processos diferentes. O princípio da segregação independente é realmente uma extensão do princípio da segregação. O princípio da segregação afirma que os dois alelos de um *locus* se separam quando os gametas são formados; o princípio da segregação independente afirma que, quando estes dois alelos se separam, sua separação é independente da separação dos alelos em *outros loci*.

Vamos ver como o princípio da segregação independente explica os resultados que Mendel obteve neste seu cruzamento di-híbrido. Cada planta tem dois alelos que codificam cada característica, então as plantas genitoras têm obrigatoriamente os genótipos *RR YY* e *rryy* (**Figura 3.9 A**). O princípio da segregação indica que os alelos para cada *locus* se separam e um alelo para cada *locus* passa para cada gameta. Os gametas produzidos pelo genitor liso e amarelo, portanto, contêm os alelos *RY*, enquanto os gametas produzidos pelo genitor rugoso e verde contêm os alelos *ry*. Estes dois tipos de gametas se unem para produzir F_3 e todos com o genótipo *Rr Yy*. Como liso é dominante sobre rugoso e amarelo é dominante sobre verde, o fenótipo de F_1 será liso e amarelo.

Quando Mendel autofertilizou as plantas F_1 para produzir F_2, os alelos para cada *locus* se separaram, com um alelo indo para cada gameta. Esse evento é aquele em que o princípio da segregação independente se torna importante. Cada par de alelos pode se separar de duas formas: (1) *R* se separa com *Y* e *r* se separa com *y*, para produzir os gametas *RY* e *ry* ou (2) *R* se separa com *y* e *r* se separa com *Y*, para produzir os gametas *Ry* e *rY*. O princípio da segregação independente nos diz que os alelos em cada *locus* se separam de forma independente; assim, os dois tipos de separação ocorrem igualmente e todos os quatro tipos de gametas (*RY, ry, Ry* e *rY*) são produzidos em proporções iguais (**Figura 3.9 B**). Quando esses quatro tipos de gametas são combinados para produzir a geração F_2, os descendentes são 9/16 lisos e amarelos, 3/16 rugosos e amarelos, 3/16 lisos e verdes e 1/16 rugoso e verde, resultando em uma razão fenotípica 9:3:3:1 (**Figura 3.9 C**).

O princípio da segregação independente e a meiose

Uma qualificação importante do princípio da segregação independente é que ela se aplica a características codificadas pelos *loci* localizados em diferentes cromossomos, porque, como o princípio da segregação, ela se baseia inteiramente no comportamento dos cromossomos na meiose. Cada par de cromossomos homólogos se separa independentemente de todos os outros pares na anáfase I da meiose (**Figura 3.10**); então os genes localizados em diferentes pares dos homólogos se segregam de forma independente. Os genes que estão localizados no mesmo cromossomo viajam juntos durante a anáfase I da meiose e chegarão ao mesmo destino – dentro do mesmo gameta (exceto se ocorrer *crossing over*). Entretanto, os genes localizados no mesmo cromossomo não se separam independentemente (exceto se estiverem distantes o suficiente para que o *crossing over* ocorra em cada divisão meiótica, como será discutido com mais detalhes no Capítulo 7).

Conceitos

O princípio da segregação independente afirma que os genes que codificam diferentes características se separam independentemente um do outro quando os gametas são formados, graças à separação independente dos pares homólogos dos cromossomos na meiose. Os genes localizados próximos no mesmo cromossomo, entretanto, não se separam de forma independente.

Checagem dos conceitos 6

Como os princípios da segregação e da segregação independente estão relacionados e quais são suas diferenças?

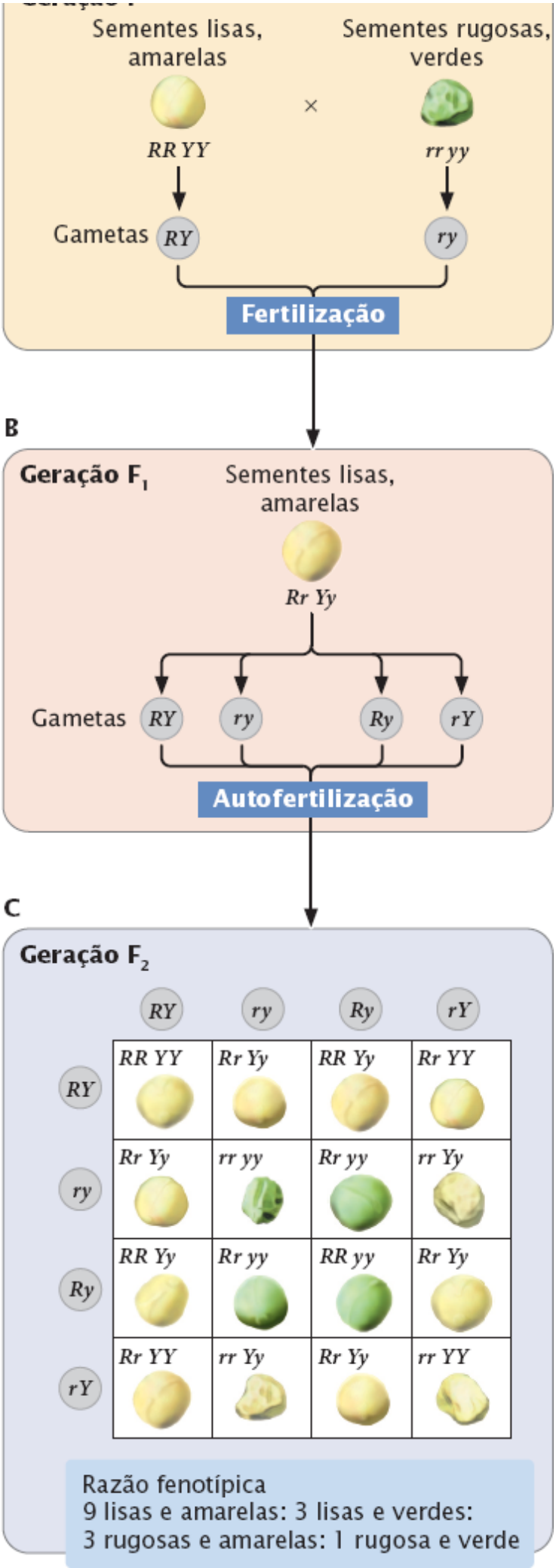
Experimento

Pergunta: Os alelos que codificam diferentes traços se separam de forma independente?

Métodos

A

Geração P


Sementes lisas, amarelas
Sementes rugosas, verdes
×
RR YY
rr yy
Gametas
RY
ry
Fertilização
B
Geração F1
Sementes lisas, amarelas
Rr Yy
Gametas
RY
ry
Ry
rY
Autofertilização
C
Resultados
Geração F2
RY
ry
Ry
rY
RR YY
Rr Yy
RR Yy
Rr YY
Rr Yy
rr yy
Rr yy
rr Yy
RR Yy
Rr yy
RR yy
Rr Yy
Rr YY
rr Yy
Rr Yy
rr YY
Razão fenotípica
9 lisas e amarelas: 3 lisas e verdes:
3 rugosas e amarelas: 1 rugosa e verde

Conclusão: O alelo que codifica a cor se separou independentemente do alelo que codifica o formato, produzindo uma razão 9:3:3:1 na geração F_2.

Figura 3.9 Os cruzamentos di-híbridos de Mendel revelaram o Princípio da Segregação Independente.

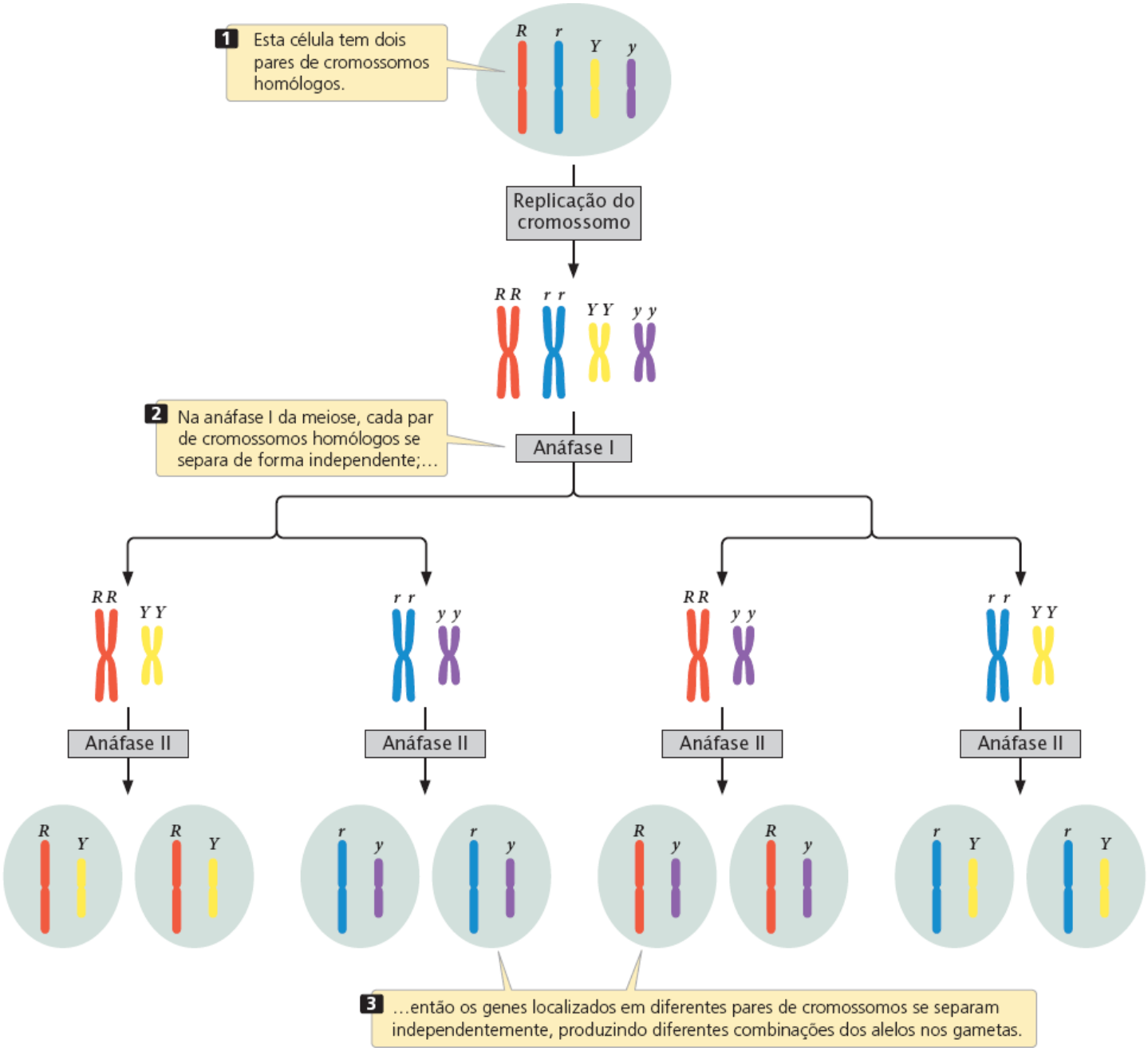


Figura 3.10 O Princípio da Segregação Independente resulta da separação independente dos cromossomos na anáfase I da meiose.

Como aplicar a probabilidade e o diagrama ramificado para os cruzamentos di-híbridos

Quando os genes em dois *loci* se separam de forma independente, um cruzamento di-híbrido pode ser compreendido como dois cruzamentos mono-híbridos. Vamos examinar o cruzamento di-híbrido de Mendel (*Rr Yy* × *Rr Yy*) ao considerar cada característica separadamente (**Figura 3.11 A**). Se considerarmos apenas o formato das sementes, o cruzamento seria *Rr* × *Rr*, o que gera uma razão fenotípica 3:1 (1/4 lisa e 1/4 rugosa dos descendentes, ver Quadro 3.5). A seguir, considere outra característica, a cor da semente. O cruzamento era *Yy* × *Yy*, que produz uma razão fenotípica 3:1 (3/4 sementes amarelas e 1/4 sementes verdes dos descendentes).

Agora podemos combinar estas razões mono-híbridas ao usar a regra de multiplicação para obter a proporção da prole com diferentes combinações de formato e cor da semente. A proporção da prole com sementes lisas e amarelas é 3/4 (a probabilidade

de ser lisa) × 3/4 (a probabilidade de ser amarela) = 9/16. A proporção da prole com sementes lisas e verdes é 3/4 × 1/4 = 3/16; a proporção da prole com sementes rugosas e amarelas é 1/4 × 3/4 = 3/16 e a proporção com sementes rugosas e verdes é 1/4 × 1/4 = 1/16.

Os diagramas ramificados são uma forma conveniente de organizar todas as combinações de características (**Figura 3.11 B**). Na primeira coluna, liste as proporções dos fenótipos para uma característica (aqui 3/4 para lisa e 1/4 para rugosa). Na segunda coluna, liste as proporções de fenótipos para a segunda característica (3/4 amarelas e 1/4 verde) duas vezes, próxima a cada um dos fenótipos na primeira coluna: coloque 3/4 amarelas e 1/4 verde próximas do fenótipo semente lisa e novamente perto do fenótipo semente rugosa. Desenhe linhas entre os fenótipos na primeira coluna e cada um dos fenótipos na segunda coluna. Agora siga cada ramificação do diagrama, multiplicando as probabilidades para cada traço ao longo da ramificação. Um ramo leva de lisa a amarela, gerando descendentes com sementes lisas e amarelas. Outro ramo leva de lisa a verde, gerando descendentes com sementes lisas e verdes, e assim em diante. Calculamos a probabilidade da prole com uma combinação particular de traços ao usar a regra de multiplicação: a probabilidade de sementes lisas (3/4) e amarelas (3/4) é 3/4 × 3/4 = 9/16. A vantagem do diagrama ramificado é que ele ajuda a manter o traço de todas as potenciais combinações que podem surgir na geração. Ele pode ser usado para determinar as proporções fenotípicas ou genotípicas para várias características.

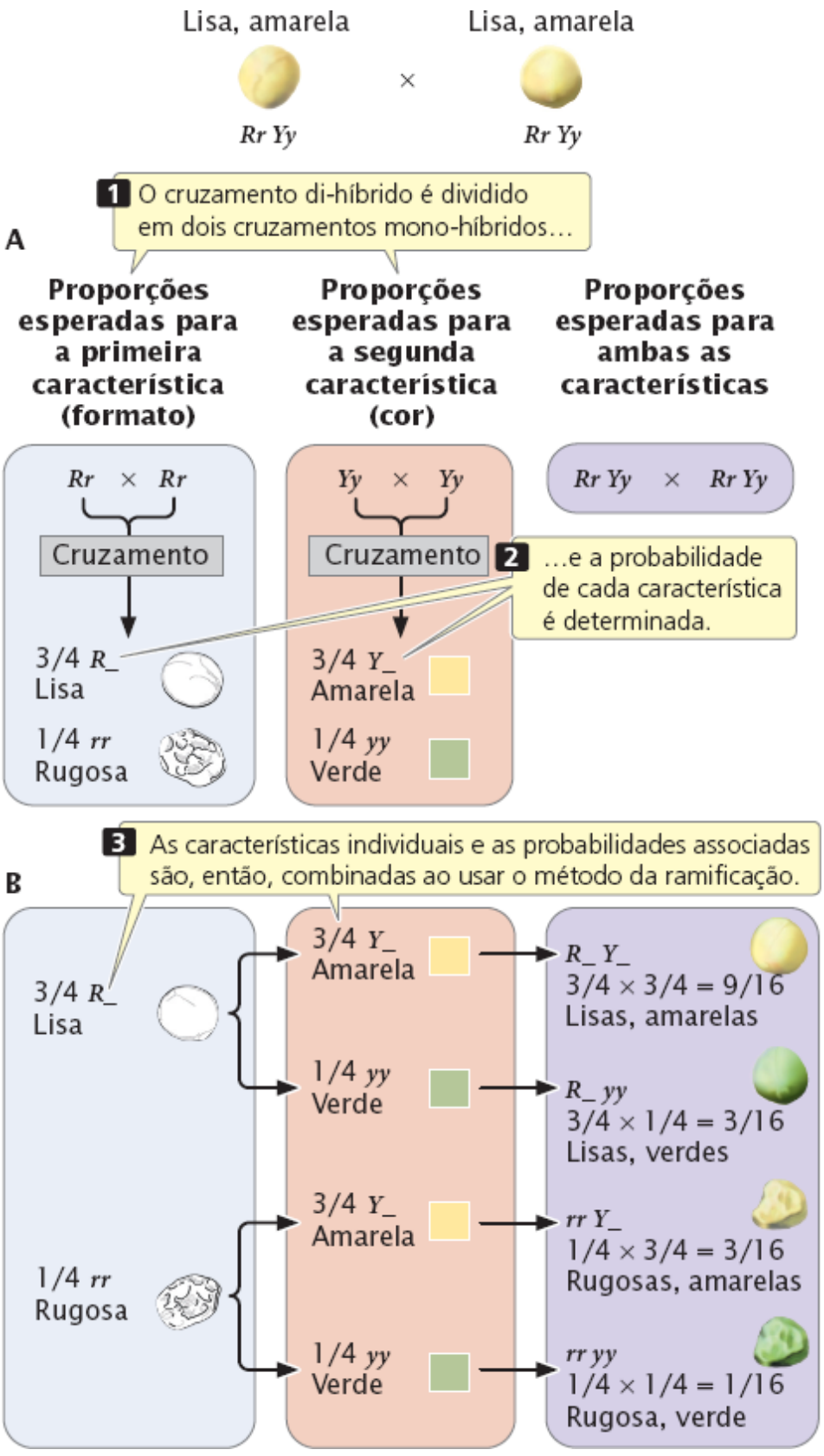


Figura 3.11 Um diagrama ramificado pode ser usado para determinar os fenótipos e proporções esperadas dos descendentes a partir de um cruzamento di-híbrido (*Rr Yy* × *Rr Yy*).

Usar a probabilidade é muito mais rápido do que usar Punnett para cruzamentos que incluem múltiplos *loci*. As razões genotípicas e fenotípicas podem ser trabalhadas de forma mais rápida ao se combinar com a regra de multiplicação as razões simples nos Quadros 3.5 e 3.6. O método de probabilidade é particularmente eficiente se precisamos da probabilidade de apenas um fenótipo ou genótipo específico na geração de um cruzamento. Suponha que precisamos saber a probabilidade de obter o genótipo *Rr yy* em F_2 do cruzamento di-híbrido na Figura 3.9. A probabilidade de obter o genótipo *Rr* em um cruzamento de *Rr* × *Rr* é de 1/2 e de obter uma geração *yy* em um cruzamento de *Yy* × *Yy* é 1/4 (ver Quadro 3.6). Usando a regra de multiplicação, encontramos a probabilidade de *Rr yy* como sendo 1/2 × 1/4 = 1/8.

Para ilustrar a vantagem do método de probabilidade, considere o cruzamento *Aa Bb cc Dd Ee* × *Aa Bb Cc dd Ee*. Suponha

que queremos saber a probabilidade de obter uma geração com o genótipo *aa bb cc dd ee*. Se usarmos um quadrado de Punnett para determinar esta probabilidade, podemos levar meses para achar a solução. Entretanto, podemos descobrir rapidamente a probabilidade de obter este genótipo ao quebrar este cruzamento em uma série de cruzamentos de um único *locus*:

Cruzamento da prole	Genótipo	Probabilidade
Aa × *Aa*	*aa*	1/4
Bb × *Bb*	*bb*	1/4
cc × *Cc*	*cc*	1/2
Dd × *dd*	*dd*	1/2
Ee × *Ee*	*ee*	1/4

Agora é fácil obter a probabilidade de uma geração a partir deste cruzamento com o genótipo *aa bb cc dd ee*, usando a regra de multiplicação: 1/4 × 1/4 × 1/2 × 1/2 × 1/4 = 1/256. Este cálculo considera que todos os genes nestes cinco *loci* se separam independentemente.

Conceitos

Um cruzamento incluindo várias características pode ser decomposto em cruzamentos de um só *locus* e usando a regra de multiplicação para determinar as proporções de combinações de características (desde que os genes se separem independentemente).

O cruzamento-teste di-híbrido

Vamos praticar usando o diagrama ramificado para determinar os tipos e proporções de fenótipos em um cruzamento-teste di-híbrido entre as plantas lisas e amarelas de F_1 (*Rr Yy*), obtidas por Mendel no seu cruzamento di-híbrido, e as plantas de sementes rugosas e verdes (*rr yy*), como descrito na **Figura 3.12**. Decomponha o cruzamento em uma série de cruzamentos de *locus* único. O cruzamento *Rr* × *rr* gera 1/2 da prole de sementes lisas (*Rr*) e 1/2 da prole de sementes rugosas (*rr*). O cruzamento *Yy yy* gera 1/2 da prole de sementes amarelas (*Yy*) e 1/2 da prole de sementes verdes (*yy*). Usando a regra de multiplicação, encontramos a proporção da prole com sementes lisas e amarelas como 1/2 (a probabilidade da característica lisa) × 1/2 (a probabilidade da característica amarela) = 1/4. Quatro combinações de traços com as seguintes proporções aparecem na prole: 1/4 *Rr Yy*, lisa e amarela; 1/4 *Rr yy*, lisa e verde; 1/4 *rr Yy*, rugosa e amarela e 1/4 *rr yy*, rugosa e verde.

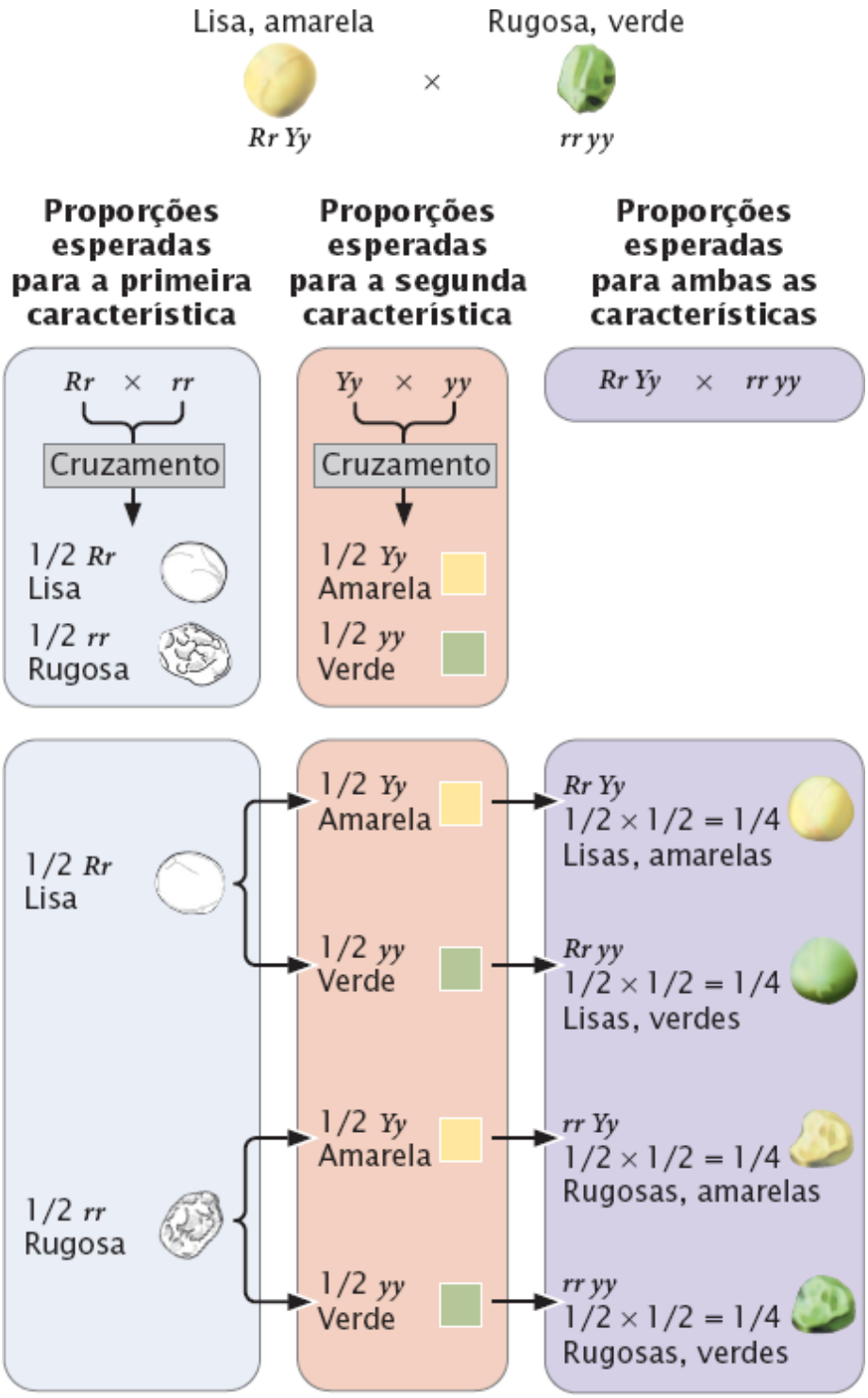


Figura 3.12 Um diagrama ramificado pode ser usado para determinar os fenótipos e as proporções esperados da prole a partir de um cruzamento-teste di-híbrido (*Rr Yy* **×** ***rr yy*****).**

Problema desenvolvido

Os princípios da segregação e segregação independente são importantes não apenas porque eles explicam como a hereditariedade funciona, mas também porque eles fornecem os meios para prever o desfecho dos cruzamentos genéticos. Este poder de prever torna a genética uma ferramenta poderosa na agricultura e em outros campos, e a capacidade de aplicar os princípios da hereditariedade é uma habilidade importante para todos os estudantes de genética. A prática com os problemas de genética é essencial para dominar os princípios básicos da hereditariedade; nenhuma leitura e memorização podem substituir a experiência obtida ao solucionar problemas específicos na genética.

Você pode achar que os problemas de genética são difíceis se você estiver inseguro sobre onde começar ou como organizar uma solução para o problema. Na genética, cada problema é diferente, então não pode ser aplicada uma série comum de etapas para todos. A lógica e o bom senso devem ser usados para analisar um problema e chegar à solução. Contudo, algumas etapas podem facilitar o processo e solucionar o problema a seguir vai ilustrar estas etapas.

Nos camundongos, a cor preta do pelo (*B*) é dominante sobre a marrom (*b*) e um padrão uniforme (*S*) é dominante sobre

o padrão manchado de branco (*s*). A cor e as manchas são controladas por genes que se separam de forma independente. Um camundongo homozigoto preto, manchado é cruzado com um camundongo homozigoto marrom, sem mancha. Todos os camundongos em F_1 são pretos e sem manchas. Um cruzamento-teste é feito ao cruzar os camundongos F_1 com camundongos marrons, com manchas.

a. Indique os genótipos dos genitores e dos camundongos de F_1.
b. Indique os genótipos e fenótipos, junto com seus padrões esperados, da prole esperada no cruzamento-teste.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

Primeiro, determine que pergunta ou perguntas o problema tem. Ele pede os genótipos ou as razões fenotípicas ou genotípicas? Este problema pede que você forneça os *genótipos* dos genitores e de F_1, os *genótipos* e *fenótipos esperados* da geração do cruzamento-teste e suas *proporções esperadas*.

Quais informações são fornecidas para solucionar o problema?

A seguir, determine quais informações fornecidas serão necessárias para solucionar o problema. Este problema fornece informações importantes sobre as relações de dominância das características e os genes que codificam os traços:

- Preto é dominante sobre o marrom
- Sem mancha é dominante sobre manchado de branco
- Os genes para as duas características se separam de forma independente
- Os símbolos para os diferentes alelos: *B* para preto, *b* para marrom, *S* para sem mancha e *s* para manchado.

É útil escrever os símbolos no início da solução:

B – preto *S* – sem mancha
b – marrom *s* – manchado de branco

A seguir, escreva os cruzamentos fornecidos no problema.

P	Homozigoto preto, manchado	×	Homozigoto marrom, sem mancha
	↓		
F_1	Preto, sem mancha		
Cruzamento-teste	Preto, sem mancha	×	Marrom, manchado

Para a solução deste problema, revise:

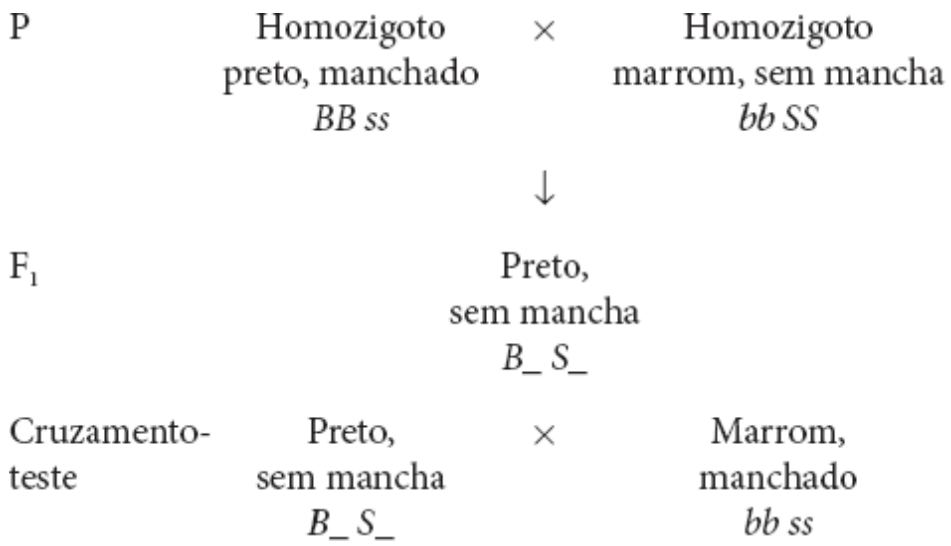
Se você precisa de ajuda para solucionar este problema, revise as seções do capítulo que discutem as informações relevantes. Para este problema, revise as Seções 3.2 e 3.3.

Etapas para a solução

Etapa 1. Escrever todas as informações genéticas que podem ser determinadas a partir dos fenótipos sozinhos.

A partir dos fenótipos e da afirmação de que são homozigotos, você sabe que a geração P de camundongos tem de ser *BB ss* e *bb SS*. Os camundongos de F_1 são pretos e sem manchas, ambos os traços dominantes; então os camundongos F_1 têm pelo menos um alelo preto (B) e um alelo sem mancha (S). Neste momento, você não tem certeza sobre outros alelos; então represente o genótipo de F_1 como *B_ S_*, em que _ significa que qualquer alelo é possível. Os camundongos marrons e manchados no cruzamento-teste devem ser *bb ss*, porque marrom e manchado são traços recessivos que serão expressos apenas se os dois alelos recessivos estiverem presentes. Registre estes genótipos nos cruzamentos que você escreveu na

etapa 2:

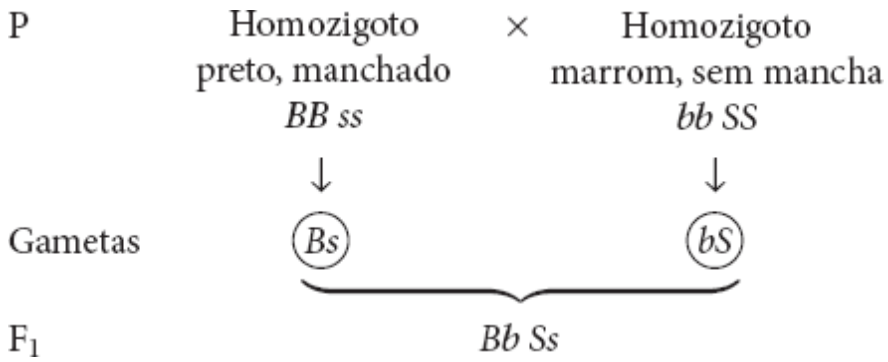


Etapa 2. Dividir o problema em partes menores.

Primeiro, determine o genótipo de F_1. Após determinar este genótipo, você pode prever os resultados do cruzamento-teste e determinar os genótipos e fenótipos da prole a partir do cruzamento-teste. Segundo, como este cruzamento inclui dois *loci* que se separam de forma independente, pode ser conveniente reduzir em dois cruzamentos de um *locus*: um para a cor do pelo e outro para a mancha. Terceiro, use um diagrama ramificado para determinar a proporção de descendentes no cruzamento-teste com diferentes combinações de dois traços.

Etapa 3. Trabalhar diferentes partes do problema.

Inicie ao determinar o genótipo da geração F_1. A primeira lei de Mendel indica que os dois alelos em um *locus* se separam, cada um indo para um gameta. Assim, os gametas produzidos pelos genitores pretos, manchados, têm *B s* e os gametas produzidos por genitores marrom, sem manchas, têm *b S*, que se combinam para produzir a prole F com o genótipo *Bb Ss*:



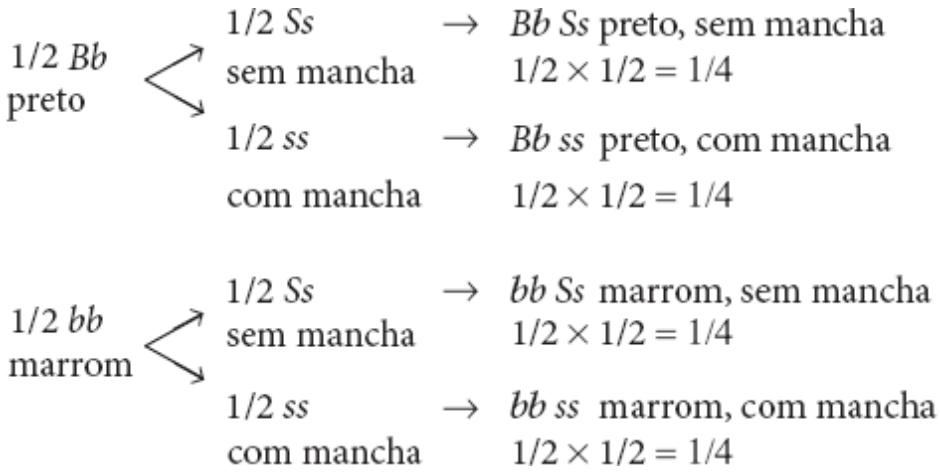
Use o genótipo F_1 para fazer o cruzamento-teste (*Bb Ss* × *bb ss*), dividindo-o em dois cruzamentos de um único *locus*. Primeiro, considere o cruzamento para a cor do pelo: *Bb* × *bb*. Qualquer cruzamento entre um genótipo recessivo heterozigoto e homozigoto produz uma proporção fenotípica 1:1 da prole (ver Quadro 3.5):

Bb × *bb*
↓
1/2 *Bb* preto
1/2 *bb* marrom

A seguir, faça o cruzamento para a característica de mancha: *Ss* × *ss*. Este cruzamento também é feito entre um genótipo recessivo heterozigoto e homozigoto e produzirá prole ½ sem mancha (*Ss*) e ½ manchada (*ss*) (ver Quadro 3.5).

Ss × *ss*
↓
1/2 *Ss* sem mancha
1/2 *ss* manchada

Finalmente, determine as proporções da prole com as combinações destas características ao usar o diagrama ramificado.



Etapa 4. Verificar todo o trabalho. Como última etapa, leia novamente o problema, verifique se suas respostas são consistentes com as informações fornecidas. Você usou os genótipos *BB ss* e *bb SS* na geração P. Estes genótipos codificam os fenótipos indicados no problema?

Os fenótipos da geração F_1 são consistentes com os genótipos que você indicou? As respostas são consistentes com as informações.

Agora que solucionamos juntos o problema de genética, resolva o **Problema 33** no final deste capítulo.

3.4 As razões observadas da prole podem desviar-se das razões esperadas ao acaso

Quando dois organismos individuais de genótipo conhecido se cruzam, esperamos certas proporções de genótipos e fenótipos nos descendentes; estas proporções se baseiam nos princípios mendelianos de segregação, segregação independente e dominância. As razões dos genótipos e fenótipos *realmente* observadas entre os descendentes, entretanto, podem desviar dessas expectativas.

Por exemplo, nas baratas-alemãs, a cor marrom do corpo (*Y*) é dominante sobre a cor amarela (*y*). Se cruzarmos uma barata marrom, heterozigota (*Yy*) com uma barata amarela (*yy*), esperamos uma proporção 1:1 de marrom (*Yy*) e amarela (*yy*) na prole. Entre 40 descendentes, portanto esperamos 20 descendentes marrons e 20 amarelos. Entretanto, os números observados podem desviar dos valores esperados; podemos de fato ver 22 marrons e 18 amarelos na prole.

O acaso tem um papel crítico nos cruzamentos genéticos, assim como jogar uma moeda. Quando você joga uma moeda, espera uma razão 1:1 – 1/2 cara e 1/2 coroa. Se você jogar a moeda 1.000 vezes, a razão de caras e coroas obtida provavelmente será muito perto da razão de 1:1. Entretanto, se você jogar a moeda 10 vezes, a razão cara e coroa pode ser bem diferente de 1:1. Você poderia obter 6 caras e 4 coroas ou 3 caras e 7 coroas, apenas por acaso. Você pode obter até 10 caras e 0 coroas. A mesma coisa acontece nos cruzamentos genéticos. Podemos esperar 20 baratas marrons e 20 amarelas, mas *poderia surgir* uma prole de 22 marrons e 18 amarelas como resultado do acaso.

Teste qui-quadrado de adequação de ajuste

Se você esperava uma razão 1:1 de baratas marrons e amarelas, mas o cruzamento produziu 22 marrons e 18 amarelas, é provável que você não ficasse muito surpreso mesmo se não fosse uma razão perfeita 1:1. Neste caso, parece razoável considerar que o acaso produziu um desvio entre os resultados esperados e os resultados observados. Mas, se você observou 25 marrons e 15 amarelas, ainda aceitaria que isto representa uma razão 1:1? Alguma coisa além do acaso pode ter causado o desvio. Talvez a herança desta característica seja mais complicada do que pensamos ou talvez alguns dos descendentes amarelos morreram antes de serem contados. Está claro que precisamos de uma forma para avaliar a probabilidade de o acaso ser responsável pelo desvio entre os números observados e os números esperados.

Para avaliar o papel do acaso na produção de desvios entre os números observados e esperados, um teste estatístico chamado de **teste qui-quadrado de adequação de ajuste** é usado. Ele fornece informações sobre como os valores observados se ajustam aos valores esperados. Antes de aprendermos como calcular o qui-quadrado, é importante compreender o que este teste indica ou não sobre um cruzamento genético.

O teste qui-quadrado não nos diz se um cruzamento genético foi feito corretamente, se os resultados são corretos ou se escolhemos a explicação genética correta para os resultados. O que ele indica é a *probabilidade* de o acaso ser responsável pela diferença entre os valores observados e esperados. Em outras palavras, ele indica a probabilidade de que o acaso sozinho possa produzir o desvio entre os valores esperados e os valores observados.

Se esperarmos uma prole de 20 marrons e 20 amarelos a partir de um cruzamento genético, o teste qui-quadrado fornece a probabilidade de que podermos observar uma geração de 25 marrons e 15 amarelos simplesmente por causa dos desvios do acaso da proporção esperada de 20:20. Esta hipótese de que o acaso sozinho é responsável por qualquer desvio entre os valores observados e esperados, é chamado em alguns casos de *hipótese nula*. A estatística, como o teste qui-quadrado, não pode provar que a hipótese nula esteja correta, mas pode nos ajudar a decidir se devemos rejeitá-la ou não. Quando a probabilidade calculada a partir do teste qui-quadrado é elevada, consideramos que o acaso sozinho produziu a diferença e não rejeitamos a hipótese nula. Quando a probabilidade é baixa, consideramos que algum fator além do acaso – algum fator importante – produziu o desvio. Por exemplo, a taxa de mortalidade das baratas amarelas pode ser maior que a das baratas marrons. Quando a probabilidade de o acaso ter induzido o desvio é baixa, rejeitamos a hipótese nula (a hipótese nula é falsa).

Para usar o teste qui-quadrado de adequação de ajuste, primeiro determinamos os resultados esperados. O teste qui-quadrado deve sempre ser aplicado a *números* da prole, não a proporções ou a porcentagens. Vamos considerar um *locus* para a cor do pelo nos gatos domésticos, que têm a cor preta (*B*) dominante sobre a cor cinza (*b*). Se cruzarmos dois gatos pretos heterozigotos (*Bb* × *Bb*), esperaríamos uma proporção 3:1 de filhotes pretos e cinzas. Uma série destes cruzamentos gera um total de 50 gatinhos – 30 pretos e 20 cinza. Estes números são nossos valores *observados*. Podemos obter os números esperados ao multiplicar as proporções *esperadas* pelo número total de descendentes observados. Neste caso, o número esperado de filhotes pretos é 3/4 × 50 = 37,5 e o número esperado de filhotes cinza é 1/4 × 50 = 12,5. O valor do qui-quadrado (x^2) é calculado ao usar a seguinte fórmula:

$$\chi^2 = \Sigma \frac{(\text{observados} - \text{esperados})^2}{\text{esperados}}$$

na qual Σ significa a soma. Calculamos a soma de todas as diferenças ao quadrado entre os valores observados e esperados e dividimos pelos valores esperados. Para calcular o valor do qui-quadrado para nossos filhotes pretos e cinzas, primeiro subtraímos o número de filhotes pretos *esperados* do número de filhotes pretos *observados* (30 – 37,5 = –7,5) e obtemos o quadrado deste valor: $–7,5^2 = 56,25$. Então dividimos este resultado pelo número esperado de filhotes pretos, 56,25/37,5 = 1,5. Repetimos os cálculos sobre o número de filhotes cinza esperados: $(20 – 12,5)^2/12,5 = 4,5$. Para obter o valor geral do qui-quadrado, somamos os valores $(\text{observados} – \text{esperados})^2/\text{esperados}$: 1,5 + 4,5 = 6,0.

O próximo passo é determinar a probabilidade associada a este valor de qui-quadrado calculado, que é a probabilidade de que o desvio entre os resultados observados e esperados poderia ser aleatória. Precisamos comparar o valor do qui-quadrado calculado (6,0) com os valores teóricos que têm os mesmos graus de liberdade em um quadro de qui-quadrado. Os graus de liberdade representam as várias formas nas quais as classes esperadas são livres para variar. Para o teste qui-quadrado de adequação de ajuste, os graus de liberdade são iguais a *n* – 1, no qual *n* é o número de diferentes fenótipos esperados. Aqui, perdemos um grau de liberdade porque o número total da geração esperada deve ser igual ao número da geração observada. Em nosso exemplo, existem dois fenótipos esperados (preto e cinza), então *n* = 2 e o grau de liberdade é 2 – 1 = 1.

Agora que temos nosso valor de qui-quadrado calculado, e descobrimos os graus de liberdade associados, estamos prontos para ter a probabilidade a partir de um quadro de qui-quadrado (**Quadro 3.7**). Os graus de liberdade são fornecidos na coluna à esquerda do quadro e as probabilidades são fornecidas no topo; dentro do corpo do quadro estão os valores de qui-quadrado associados a estas probabilidades. Primeiro, encontre a coluna para os graus de liberdade apropriados, para nosso exemplo com 1 grau de liberdade, é a primeira coluna do quadro. Descubra onde nosso valor de qui-quadrado calculado (6,0) está entre os valores teóricos nesta coluna. Os valores teóricos do qui-quadrado aumentam da esquerda para direita e as probabilidades reduzem da esquerda para direita. Nosso valor de qui-quadrado de 6,0 fica entre o valor de 5,024, associado a uma probabilidade de 0,025 e o valor de 6,635, associado a uma probabilidade de 0,01.

Quadro 3.7 Valores críticos da distribuição de x^2.

					P				
gl	**0,995**	**0,975**	**0,9**	**0,5**	**0,1**	**0,05***	**0,025**	**0,01**	**0,005**
1	0,000	0,000	0,016	0,455	2,706	3,841	5,024	6,635	7,879
2	0,010	0,051	0,211	1,386	4,605	5,991	7,378	9,210	10,597

3	0,072	0,216	0,584	2,366	6,251	7,815	9,348	11,345	12,838
4	0,207	0,484	1,064	3,357	7,779	9,488	11,143	13,277	14,860
5	0,412	0,831	1,610	4,351	9,236	11,070	12,832	15,086	16,750
6	0,676	1,237	2,204	5,348	10,645	12,592	14,449	16,812	18,548
7	0,989	1,690	2,833	6,346	12,017	14,067	16,013	18,475	20,278
8	1,344	2,180	3,490	7,344	13,362	15,507	17,535	20,090	21,955
9	1,735	2,700	4,168	8,343	14,684	16,919	19,023	21,666	23,589
10	2,156	3,247	4,865	9,342	15,987	18,307	20,483	23,209	25,188
11	2,603	3,816	5,578	10,341	17,275	19,675	21,920	24,725	26,757
12	3,074	4,404	6,304	11,340	18,549	21,026	23,337	26,217	28,300
13	3,565	5,009	7,042	12,340	19,812	22,362	24,736	27,688	29,819
14	4,075	5,629	7,790	13,339	21,064	23,685	26,119	29,141	31,319
15	4,601	6,262	8,547	14,339	22,307	24,996	27,488	30,578	32,801

P, probabilidade; gl, graus de liberdade.

*A maior parte dos cientistas considera que, quando $P < 0,05$, existe uma diferença significativa entre os valores observados e esperados em um teste qui-quadrado.

Assim, a probabilidade associada a nosso valor de qui-quadrado é menor que 0,025 e maior que 0,01. Então existe uma probabilidade menor que 2,5% de que o acaso seja responsável pelo desvio que observamos entre os números esperados e observados dos filhotes pretos e cinza.

Muitos cientistas usam o nível de probabilidade de 0,05 como valor de corte: se a probabilidade do acaso responsável pelo desvio for maior ou igual a 0,05, eles aceitam que este possa ser responsável pelo desvio entre os valores observados e esperados. Quando a probabilidade é menor que 0,05, os cientistas consideram que o acaso não é responsável e existe uma diferença significativa. A expressão *diferença significativa* indica que um fator além do acaso é responsável pelos valores observados diferentes dos valores esperados. Em relação aos filhotes, talvez um dos genótipos tivesse uma maior taxa de mortalidade antes dos descendentes serem contados ou talvez outros fatores genéticos tenham desviado as razões observadas.

Ao escolher 0,05 como valor de corte, os cientistas concordam que o acaso é responsável pelos desvios entre os valores observados e esperados exceto se existir uma forte evidência do contrário. Tenha em mente que, mesmo se tivermos uma probabilidade de, por exemplo, 0,01, ainda existe uma probabilidade de 1% de que o desvio entre os números observados esperado não seja nada além do acaso. O cálculo do valor de qui-quadrado está ilustrado na **Figura 3.13**. Resolva o Problema 38

Conceitos

As diferenças entre as proporções observadas e esperadas podem surgir ao acaso. O teste qui-quadrado de adequação de ajuste pode ser usado para avaliar se é provável que o acaso ou outro fator importante seja responsável pelos desvios entre os números observados e esperados.

Checagem dos conceitos 7

É feito um teste qui-quadrado comparando os descendentes observados e esperados e a probabilidade associada ao valor de qui-quadrado calculado é 0,72. O que esta probabilidade representa?

a. Probabilidade de que foram obtidos os resultados corretos.
b. Probabilidade de obter os números observados.
c. Probabilidade de que a diferença entre os números observados e esperados seja importante.
d. Probabilidade de que a diferença entre os números observados e esperados seja o acaso.

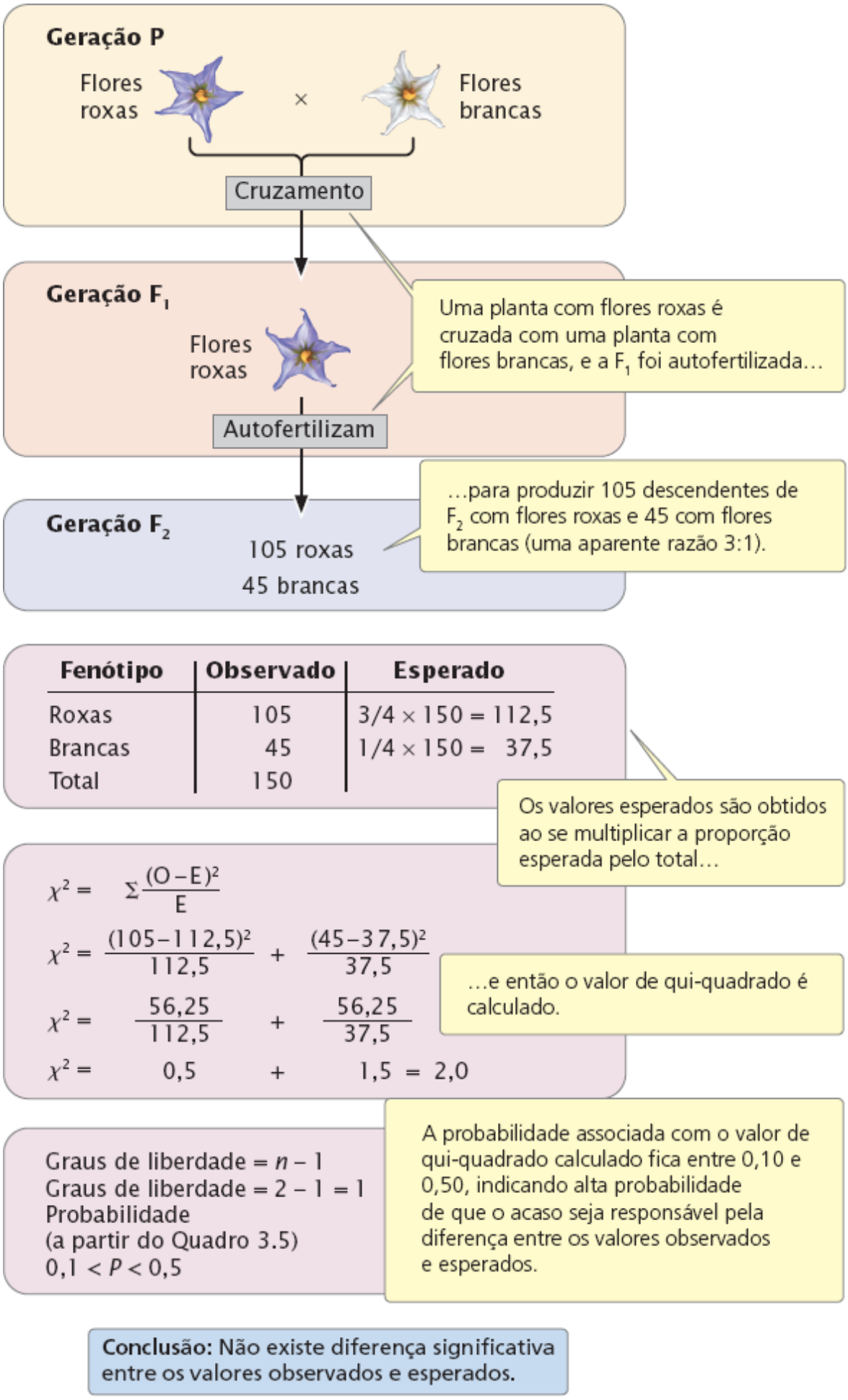


Figura 3.13 **Um teste qui-quadrado é usado para determinar a probabilidade de o acaso ser responsável pela diferença entre os valores observados e esperados.**

Resumo dos conceitos

- Gregor Mendel descobriu os princípios básicos da hereditariedade. Seu sucesso pode ser atribuído a sua escolha da ervilha como organismo experimental, o uso de características com fenótipos de fácil distinção, sua abordagem experimental, o uso de matemática para interpretar seus resultados e a atenção cuidadosa aos detalhes
- Os genes são fatores hereditários que determinam uma característica. As formas alternativas de um gene são chamadas de alelos. Os alelos estão localizados em local específico, um *locus*, em um cromossomo, e o conjunto de genes que um organismo individual possui é o seu genótipo. O fenótipo é a manifestação ou aparecimento de uma característica e pode se referir a uma característica física, bioquímica ou comportamental. Apenas o genótipo é herdado e não o fenótipo
- O princípio da segregação afirma que um organismo diploide tem dois alelos que codificam um traço e que estes dois alelos se separam em proporções iguais quando os gametas são formados
- O conceito de dominância afirma que quando dois alelos diferentes estão presentes em um heterozigoto, apenas o traço codificado por um deles – o alelo "dominante" – é observado no fenótipo. O outro alelo é recessivo
- Os dois alelos de um genótipo estão localizados nos cromossomos homólogos. A separação dos cromossomos homólogos na anáfase I da meiose leva à segregação dos alelos
- A probabilidade é o cálculo da ocorrência de um evento particular acontecer. A regra de multiplicação afirma que a probabilidade de dois ou mais eventos independentes ocorrerem juntos é calculada ao multiplicar as probabilidades dos eventos independentes. A regra de adição afirma que a probabilidade de um de dois ou mais eventos mutuamente excludentes ocorrer é calculada ao somar as probabilidades dos eventos
- A expansão binomial pode ser usada para determinar a probabilidade de um conjunto específico de eventos
- Um cruzamento-teste revela se o genótipo (homozigoto ou heterozigoto) de um organismo tem um traço dominante e consiste no cruzamento deste indivíduo com outro de genótipo recessivo homozigoto
- O princípio da segregação independente afirma que os genes que codificam diferentes características se separam de forma independente quando os gametas são formados. A segregação independente se baseia na separação aleatória dos pares homólogos dos cromossomos na anáfase I da meiose; ela ocorre quando os genes que codificam duas características estão localizados em diferentes pares de cromossomos
- As razões observadas a partir de um cruzamento genético podem desviar-se das razões esperadas graças ao acaso. O teste qui-quadrado é usado para determinar a probabilidade de que o acaso seja responsável pela diferença entre os valores observados e esperados.

Termos importantes

Alelo
Conceito de dominância
Cruzamento di-híbrido
Cruzamento mono-híbrido
Cruzamento recíproco
Cruzamento-teste
Dominante
Fenótipo
Gene
Genótipo
Geração F_1
Geração F_2
Geração P
Heterozigoto
Homozigoto
Locus
Princípio da segregação (primeira lei de Mendel)
Princípio da segregação independente (segunda lei de Mendel)
Probabilidade
Probabilidade condicional
Quadrado de Punnett
Recessivo
Regra de adição

Regra de multiplicação
Retrocruzamento
Teoria cromossômica da hereditariedade
Teste qui-quadrado de adequação de ajuste
Tipo selvagem

Respostas da Checagem dos conceitos

1. b.
2. *Locus* é uma posição no cromossomo onde estão localizadas as informações genéticas que codificam uma característica. Um alelo é uma versão de um gene que codifica um traço específico. Um genótipo é um conjunto de alelos que um organismo individual tem e fenótipo é a manifestação ou aparência de uma característica.
3. Porque os traços para ambos os alelos aparecem na prole de F_2.
4. d.
5. a.
6. O princípio da segregação e o princípio da segregação independente se referem à separação dos alelos na anáfase I da meiose. O princípio da segregação afirma que estes alelos se separam e o da segregação independente afirma que eles se separam de forma independente dos alelos em outros *loci*.
7. d.

Problemas desenvolvidos

Problema 1

Pelo curto (*S*) nos coelhos é dominante sobre o pelo longo (*s*). Foram realizados os seguintes cruzamentos, produzindo a prole a seguir. Indique todos os possíveis genótipos dos genitores em cada cruzamento.

Genitores	Prole
a. Curto × curto	4 curtos e 2 longos
b. Curto × curto	8 curtos
c. Curto × longo	12 curtos
d. Curto × longo	3 curtos e 1 longo
e. Longo × longo	2 longos

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Todos os possíveis genótipos dos genitores em cada cruzamento.

Quais informações são fornecidas para solucionar o problema?

- O pelo curto é dominante sobre o pelo longo
- Fenótipos dos genitores de cada cruzamento
- Fenótipos e número de descendentes de cada cruzamento.

Para a solução deste problema, revise:
Seção 3.2.

Etapas para a solução

Para este problema, é útil coletar o máximo de informações sobre os genótipos dos genitores se possível com base nos seus fenótipos. Podemos observar os tipos de descendentes produzidos para obter as informações ausentes.

Nota: O problema pede todos os possíveis genótipos dos genitores.

a. Curto × curto — 4 curtos e 2 longos

Como o pelo curto é dominante sobre o pelo longo, um coelho de pelo curto poderia ser *SS* ou *Ss*. Os dois descendentes de pelo longo têm de ser homozigotos (*ss*) porque o pelo longo é recessivo e aparecerá no fenótipo apenas quando os dois alelos para pelo longo estiverem presentes. Como cada genitor contribui com um dos dois alelos encontrados nos descendentes, cada genitor tem de ter o alelo *s* e, portanto, é *Ss*.

Nota: Quando ambos os genitores são heterozigotos (*Ss* × *Ss*), esperamos uma proporção 3:1 nos descendentes, mas, por acaso, eles exibiram uma proporção 4:2.

b. Curto × curto — 8 curtos

Os genitores de pelo curto podem ser *SS* ou *Ss*. Todos os oito descendentes têm pelo curto (*S_*), então é provável que pelo menos um dos genitores seja homozigoto (*SS*). Se ambos os genitores fossem heterozigotos, esperaríamos 1/4 dos descendentes de pelo longo (*ss*), mas não observamos nenhum descendente desse tipo. O outro genitor poderia ser homozigoto (*SS*) ou heterozigoto (*Ss*); desde que um genitor seja homozigoto, todos os descendentes terão pelo curto. Teoricamente é possível, embora improvável, que ambos os genitores sejam heterozigotos (*Ss* × *Ss*). Se ambos forem heterozigotos, esperaríamos que dois dos oito descendentes tivessem pelo longo. Embora não seja observado nenhum descendente de pelo longo, é possível que por acaso não tenha sido gerado nenhum coelho de pelo longo entre os oito descendentes do cruzamento.

c. Curto × longo — 12 curtos

O genitor de pelo curto pode ser *SS* ou *Ss*. O genitor de pelo longo deve ser *ss*. Se o genitor de pelo curto era heterozigoto (*Ss*), seria esperado que metade da geração tivesse pelo longo, mas não há nenhum descendente de pelo longo. Portanto, é mais provável que este genitor seja homozigoto (*SS*). É teoricamente possível, embora improvável, que o genitor seja heterozigoto e por acaso não foi gerado descendente de pelo longo.

d. Curto × longo — 3 curtos e 1 longo

Com base no seu fenótipo, o genitor de pelo curto poderia ser homozigoto (*SS*) ou heterozigoto (*Ss*), mas a prole de pelo longo nos diz que o genitor de pelo curto tem de ser heterozigoto (*Ss*). O genitor de pelo longo deve ser homozigoto (*ss*).

e. Longo × longo — 2 longos

Como o pelo longo é recessivo, ambos os genitores têm de ser homozigotos para um alelo de pelo longo (*ss*).

Problema 2

Nos gatos, o pelo preto é dominante sobre o pelo cinza. Uma gata preta cuja mãe é cinza cruza com um gato cinza. Se esta fêmea tem uma ninhada de seis filhotes, qual é a probabilidade de que três serão pretos e três serão cinza?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
A probabilidade de que, em uma ninhada de seis filhotes, três serão pretos e três serão cinza.

Quais informações são fornecidas para solucionar o problema?

- Preto é dominante sobre cinza
- A mãe da ninhada é preta e sua mãe é cinza
- O pai da ninhada é cinza.

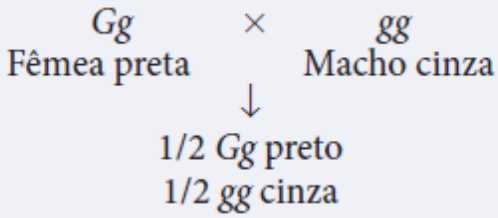
Para a solução deste problema, revise:

Seção 3.2.

Etapas para a solução

Como preto (*G*) é dominante sobre cinza (*g*), um gato preto pode ser homozigoto (*GG*) ou heterozigoto (*Gg*). A fêmea preta neste problema tem de ser heterozigota (*Gg*) porque sua mãe é cinza (*gg*) e ela herdou um dos alelos da mãe. O macho cinza é homozigoto (*gg*) porque cinza é recessivo. Desta forma, o cruzamento é:

Dica: Podemos determinar o genótipo da fêmea a partir do seu fenótipo e do fenótipo da sua genitora.



Podemos usar a expansão binomial para determinar a probabilidade de obter três filhotes pretos e três filhotes cinza em uma ninhada de seis. Neste caso, *p* é a probabilidade de um filhote ser preto e *q* é a probabilidade de um filhote ser cinza. O binômio é $(p + q)^6$ e sua expansão é:

Lembrete: A expansão binomial pode ser usada para determinar a probabilidade de diferentes combinações de traços na prole de um cruzamento.

$$(p + q)^6 = p^6 + 6p^5q + 15p^4q^2 + 20p^3q^3 + 15p^2q^4 + 6p^1q^5 + q^6$$

A probabilidade de obter três filhotes pretos e três filhotes cinza em uma ninhada de seis é indicada pelo termo $20p^3q^3$. As probabilidades de *p* e *q* são 1/2; então a probabilidade é $20(1/2)^3\ (1/2)^3 = 20/64 = 5/16$.

Dica: Ler, nas páginas 50-51, a explicação de como expandir o binômio.

Problema 3

No milho, as sementes roxas são dominantes sobre as sementes amarelas, e as sementes lisas são dominantes sobre as sementes enrugadas. Um milho com sementes roxas e lisas é cruzado com um milho com sementes amarelas e enrugadas e são obtidos os seguintes descendentes:

Roxas, lisas	112
Roxas, enrugadas	103
Amarelas, lisas	91
Amarelas, enrugadas	94

Quais são os genótipos mais prováveis dos genitores e da prole? Teste sua hipótese genética com um teste qui-quadrado.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

Os genótipos dos genitores e da prole.

Um teste qui-quadrado comparando os resultados observados e esperados.

Quais informações são fornecidas para solucionar o problema?

- As sementes roxas são dominantes sobre as sementes amarelas, e as sementes lisas são dominantes sobre as sementes enrugadas
- Os fenótipos dos genitores
- Os fenótipos e os vários descendentes diferentes do cruzamento.

Para a solução deste problema, revise:
Seções 3.3 e 3.4.

Etapas para a solução

A melhor forma para começar este problema é dividindo o cruzamento em cruzamentos simples para uma única característica (cor ou formato da semente):

P	roxa × amarela	lisa × enrugada
F_1	112 + 103 = 215 roxas 91 + 94 = 185 amarelas	112 + 91 = 203 lisas 103 + 94 = 197 enrugadas

Neste cruzamento, roxa × amarela produz aproximadamente 1/2 roxa e 1/2 amarela, uma razão 1:1. Essa razão surge, em geral, a partir de um cruzamento entre um heterozigoto e um homozigoto. Como roxa é dominante, o genitor de semente roxa tem de ser heterozigoto (*Pp*) e o genitor de semente amarela tem de ser homozigoto (*pp*). Os descendentes de semente roxa produzidos por este cruzamento serão heterozigotos (*Pp*) e os descendentes de semente amarela têm de ser homozigotos (*pp*).

Dica: Uma boa estratégia em um cruzamento envolvendo múltiplas características é analisar separadamente os resultados para cada característica.

Vamos examinar a outra característica. As características lisa × enrugada produzem 1/2 lisa e 1/2 enrugada, uma razão 1:1, então os fenótipos da prole também são produzidos por um cruzamento entre um heterozigoto (*Ff*) e um homozigoto (*ff*); os descendentes de semente lisa serão heterozigotos (*Ff)* e os descendentes de semente enrugada serão homozigotos (*ff*).

Agora combine os dois cruzamentos e use a regra de multiplicação para obter os genótipos completos e as proporções de cada genótipo:

Lembrete: A regra de multiplicação afirma que a probabilidade de dois ou mais eventos independentes ocorrerem juntos é calculada ao multiplicar suas probabilidades independentes.

P	roxa, lisa *Pp Ff*	×	amarela, enrugada *pp ff*
F_1	*Pp Ff* = 1/2 roxa	×	1/2 lisa = 1/4 roxa, lisa
	Pp ff = 1/2 roxa	×	1/2 enrugada = 1/4 roxa, enrugada
	pp Ff = 1/2 amarela	×	1/2 lisa = 1/4 amarela, lisa
	pp ff = 1/2 amarela	×	1/2 enrugada = 1/4 amarela, enrugada

Nossa explicação genética prevê que, para este cruzamento, devemos ver descendentes de sementes 1/4 roxas e lisas, 1/4 roxas e enrugadas, 1/4 amarelas e lisas, e 1/4 amarelas e enrugadas. Foram produzidos 400 descendentes, então é esperada uma estimativa de 1/4 × 400 = 100 de cada fenótipo. Estes números observados não se ajustam exatamente aos números esperados. A diferença entre o que observamos e o que esperamos poderia ser obra do acaso? Se a probabilidade de que o acaso sozinho seja responsável pela diferença entre os valores observados e esperados é alta, consideraremos que os descendentes foram produzidos na razão 1:1:1:1 prevista pelo cruzamento. Se for baixa a probabilidade de que a diferença entre os valores observados e esperados seja obra do acaso, a prole não está de fato na razão prevista e outro

fator *significativo* tem de ser responsável pelo desvio.

Os números observados e esperados são:

Fenótipo	Observado	Esperado
Roxa, lisa	112	1/4 × 400 = 100
Roxa, enrugada	103	1/4 × 400 = 100
Amarela, lisa	91	1/4 × 400 = 100
Amarela, enrugada	94	1/4 × 400 = 100

Para determinar a probabilidade de que a diferença entre os valores observados e esperados seja obra do acaso, calculamos um valor do qui-quadrado com a fórmula $x^2 = \Sigma[(\text{observado} - \text{esperado})^2/\text{esperado}]$:

Dica: Ver Figura 3.12 para ajudar a fazer um teste qui-quadrado.

$$\chi^2 = \frac{(112-100)^2}{100} + \frac{(103-100)^2}{100} + \frac{(91-100)^2}{100} + \frac{(94-100)^2}{100}$$
$$= \frac{12^2}{100} + \frac{3^2}{100} + \frac{9^2}{100} + \frac{6^2}{100}$$
$$= \frac{144}{100} + \frac{9}{100} + \frac{81}{100} - \frac{36}{100}$$
$$= 1{,}44 + 0{,}09 + 0{,}81 + 0{,}36 = 2{,}70$$

Agora que já temos o valor de qui-quadrado, é preciso determinar a probabilidade de que o acaso seja responsável por esse valor. Para obter esta probabilidade, primeiro calculamos os graus de liberdade, para os quais o teste qui-quadrado de adequação de ajuste é $n - 1$, em que n é igual ao número de classes fenotípicas esperadas. Neste caso, existem quatro classes fenotípicas esperadas, de forma que os graus de liberdade são iguais a 4 – 1 = 3. Agora, é preciso determinar o valor do qui-quadrado em um quadro de qui-quadrado (ver Quadro 3.7). Selecionamos a coluna correspondente ao terceiro grau de liberdade e olhamos ao longo dela para encontrar nosso valor de qui-quadrado calculado. O valor de qui-quadrado calculado de 2,7 está entre 2,366 (uma probabilidade de 0,5) e 6,251 (uma probabilidade de 0,1). A probabilidade (P) associada ao valor de qui-quadrado calculado é, portanto, $0{,}5 < P < 0{,}1$. Este valor é a probabilidade de que a diferença entre o que observamos e o que esperamos seja decorrente do acaso, que neste caso é relativamente alta, então é provável que o acaso seja responsável pelo desvio. Podemos concluir que os descendentes aparecem na razão 1:1:1:1 prevista pela nossa explicação genética.

Questões de compreensão

Seção 3.1

1. Por que a abordagem de Mendel para o estudo da hereditariedade teve tanto sucesso?
2. Qual a diferença entre genótipo e fenótipo?

Seção 3.2

3. O que é o princípio da segregação? Por que é importante?
4. Como os princípios de Mendel são diferentes do conceito de herança por mistura discutido no Capítulo 1?
5. O que é o conceito de dominância?
6. O que são as regras de multiplicação e adição de probabilidade e quando devem ser usadas?
7. Forneça as razões genotípicas que podem aparecer entre os descendentes de cruzamentos simples e os genótipos dos genitores que podem dar origem a cada proporção.
8. O que é a teoria cromossômica da herança. Por que ela é importante?

Seção 3.3

9. O que é o princípio da segregação independente? Como ele está relacionado com o princípio da segregação?
10. Em quais fases da mitose e meiose atuam os princípios da segregação e segregação independente?

Seção 3.4

11. Como o teste qui-quadrado de adequação de ajuste é usado para analisar os cruzamentos genéticos? O que a probabilidade associada a um valor de qui-quadrado indica sobre os resultados de um cruzamento?

Questões e problemas aplicados

Introdução

12. A herança do cabelo ruivo foi discutida na introdução deste capítulo. Em alguns momentos do passado, acreditava-se que o cabelo ruivo fosse um traço recessivo; em outros, que fosse um traço dominante. Quais características de herança o cabelo ruivo deveria exibir como um traço recessivo? Quais características seriam esperadas se ele fosse um traço dominante?

Seção 3.1

*13. Quais características de um organismo o tornariam adequado para estudos dos princípios da herança? Você pode citar alguns organismos com essas características?

Seção 3.2

14. Nos pepinos, a cor laranja (*R*) é dominante sobre a cor creme (*r*). Um pepino homozigoto para cor laranja é cruzado com um pepino homozigoto para cor creme. F_1 é intercruzada para produzir F_2.
    a. Indique os genótipos e fenótipos dos genitores, de F_1 e de F_2.
    b. Indique os genótipos e fenótipos da prole de um retrocruzamento entre F_1 e o genitor de cor laranja.
    c. Indique os genótipos e fenótipos da prole de um retrocruzamento entre F_1 e o genitor de cor creme.
15. Imagine três garotas descendentes dos nativos americanos hospi, uma com albinismo. As três garotas poderiam ser irmãs? Por que sim ou não?

16. J. W. McKay cruzou um melão-cantalupo que produziu sementes castanhas com uma planta que produziu sementes vermelhas e obteve os seguintes resultados (J. W. McKay. 1936. *Journal of Heredity* 27:110-112).

Cruzamento	F_1	F_2
Castanho ♀ × vermelho ♂	13 sementes castanhas	93 sementes castanhas, 24 vermelhas

    a. Explique a herança das sementes castanhas e vermelhas nesta planta.
    b. Indique símbolos para os alelos neste cruzamento e os genótipos para todas as plantas individuais.

*17. A cor branca (*w*) do pelo dos porquinhos-da-índia é recessiva para preta (*W*). Em 1909, W. E. Castle e J. C. Phillips transplantaram um ovário de uma fêmea preta de porquinho-da-índia em uma fêmea branca que teve seus ovários removidos. Então eles cruzaram esta fêmea branca com um macho branco. Toda a prole do cruzamento tinha cor preta (W. E. Castle and J. C. Phillips. 1909. *Science* 30:312-313).
    a. Explique os resultados deste cruzamento.
    b. Indique o genótipo dos descendentes deste cruzamento.
    c. O que esse experimento pode indicar sobre a validade das teorias de pangênese e germeplasma discutidas no Capítulo 1?

*18. Nos gatos, o tipo sanguíneo A resulta de um alelo (I^A), dominante sobre o alelo (i^B), que produz o tipo B. Não existe o tipo 0. Apresentamos a seguir os tipos sanguíneos de gatos e gatas que foram cruzados e os tipos sanguíneos dos seus filhotes. Indique os genótipos mais prováveis para os genitores de cada ninhada.

	Genitor macho	Genitora fêmea	Filhotes
a.	A	B	4 com tipo A, 3 com tipo B

b.	B	B	6 com tipo B
c.	B	A	8 com tipo A
d.	A	A	7 com tipo A, 2 com tipo B
e.	A	A	10 com tipo A
f.	A	B	4 com tipo A, 1 com tipo B

19. A **Figura 3.7** mostra os resultados de um cruzamento entre as plantas alta e baixa da ervilha.
    a. Quais fenótipos e proporções serão produzidos se for feito um retrocruzamento entre os descendentes altos F_1 e o genitor baixo?
    b. Quais fenótipos e proporções serão produzidos se for feito um retrocruzamento entre os descendentes altos F_1 e o genitor alto?
20. João tinha um gato branco chamado Sam. Quando João cruzou Sam com uma gata preta, ele obteve uma ninhada de 1/2 de gatinhos brancos e 1/2 de gatinhos pretos. Quando os gatinhos pretos cruzaram entre si, todos os seus filhotes eram pretos. Com base nesses resultados, você concluiria que o traço recessivo é a cor preta ou a cor branca do pelo do gato? Explique seu raciocínio.

*21. No carneiro, a lã brilhante resulta de um alelo (*L*) que é dominante sobre um alelo (*l*) da lã normal. Uma ovelha (fêmea adulta) com lã brilhante é cruzada com um carneiro (adulto macho) com lã normal. A ovelha então gerou um cabrito com lã normal. A partir deste único descendente, é possível determinar os genótipos dos dois genitores? Neste caso, quais são seus genótipos? Explique se não for possível.

*22. Alcaptonúria é um distúrbio metabólico na qual as pessoas produzem urina escura. Ela é oriunda de um alelo (*a*) que é recessivo para o alelo para o metabolismo normal (*A*). Sally tem o metabolismo normal, mas seu irmão tem alcaptonúria. O pai de Sally tem alcaptonúria, e a mãe tem o metabolismo normal.
    a. Indique os genótipos de Sally, sua mãe, seu pai e seu irmão.
    b. Se os pais de Sally tiverem outra criança, qual é a probabilidade de que essa criança terá alcaptonúria?
    c. Se Sally se casar com um homem com alcaptonúria, qual é a probabilidade de que seu primeiro filho tenha alcaptonúria?
23. Suponha que você esteja criando gerbos (um tipo de roedor). E observou que alguns dos seus roedores têm manchas brancas, enquanto outros não têm. Que tipo de cruzamentos você precisa fazer para determinar se as manchas brancas são causadas por um alelo recessivo ou dominante?
24. A calvície nos rat terrier americanos é recessiva para a presença de pelo. Imagine que você tenha um terrier com pelo. Como você pode identificar se este cão é homozigoto ou heterozigoto para o traço do pelo?

*25. Qual é a probabilidade de jogar um dado de seis lados e obter os seguintes números?
    a. 2.
    b. 1 ou 2.
    c. Um número par.
    d. Qualquer número exceto o seis.

*26. Qual é a probabilidade de jogar um dado de seis lados duas vezes e obter os seguintes números?
    a. 2 e 3.
    b. 6 e 6.
    c. Pelo menos um 6.
    d. Dois do mesmo número (dois 1 ou dois 2 ou dois 3 etc.).
    e. Um número par nas duas jogadas.
    f. Um número par ou pelo menos um número ímpar.

*27. Em uma família de sete crianças, qual é a probabilidade de ter os seguintes números de meninos e meninas?
    a. Todos meninos.
    b. Todas as crianças do mesmo sexo.
    c. Seis meninas e um menino.
    d. Quatro meninos e três meninas.
    e. Quatro meninas e três meninos.

28. A fenilcetonúria (PKU) é uma doença que resulta de um gene recessivo. Dois pais normais têm uma criança com PKU.
    a. Qual a probabilidade de um espermatozoide do pai ter o alelo para PKU?
    b. Qual a probabilidade de um óvulo da mãe ter o alelo para PKU?
    c. Qual a probabilidade de a próxima criança ter PKU?
    d. Qual a probabilidade de a próxima criança ser heterozigota para o gene PKU?

*29. Nas baratas alemãs, a asa curva (*cv*) é recessiva para asa normal (cv^+). Uma barata homozigota com asas normais é

cruzada com uma barata homozigota com asas curvas. F_1 é intercruzada para produzir F_2. Considere que o par de cromossomos com o *locus* para o formato da asa seja metacêntrico. Desenhe este par de cromossomos como ele apareceria nos genitores, na F_1 e em cada classe de descendentes de F_2 na metáfase I da meiose. Considere que não ocorre *crossing over*. Em cada estágio, indique a localização dos alelos para o formato da asa (*cv* e cv^+) nos cromossomos.

*30. Nos porquinhos-da-índia, o alelo para o pelo preto (*B*) é dominante sobre o alelo para pelo marrom (*b*). Um cobaio preto é cruzado com um marrom, produzindo cinco cobaios de pelo preto e seis com pelo marrom na F_1.
   a. Quantas cópias do alelo preto (*B*) estarão presentes em cada célula de um porquinho preto nos seguintes estágios: G_1, G_2, metáfase da mitose, metáfase I da meiose, metáfase II da meiose e após a segunda citocinese seguinte à meiose? Considere que não ocorre *crossing over*.
   b. Quantas cópias do alelo marrom (*B*) existem em cada célula de um cobaio marrom nos mesmos estágios listados na alternativa a? Considere que não ocorre *crossing over*.

Seção 3.3

31. Nas melancias, a fruta amarga (*B*) é dominante sobre a fruta doce (*b*) e a presença de manchas amarelas (*S*) é dominante sobre a ausência de manchas (*s*). Os genes para estas duas características se separam de forma independente. Uma planta homozigota com fruta amarga e manchas amarelas é cruzada com uma planta homozigota com fruta doce e sem manchas. F_1 é intercruzada para produzir F_2.
   a. Quais são as razões fenotípicas em F_2?
   b. Se uma planta F_1 fizer retrocruzamento com a genitora de frutas amargas e manchas amarelas, quais são os fenótipos e as proporções esperados na prole?
   c. Se uma planta F_1 fizer retrocruzamento com a genitora de frutas doces e sem manchas, quais são os fenótipos e as proporções esperados na prole?
32. A **Figura 3.9** mostra os resultados de um cruzamento di-híbrido envolvendo o formato e cor da semente.
   a. Qual é a proporção de descendentes F_2 lisos e amarelos a partir deste cruzamento em ambos os *loci*?
   b. Qual é a proporção de descendentes F_2 lisos e amarelos a partir deste cruzamento em pelo menos um *locus*?

*33. Nos gatos, as orelhas curvas resultam de um alelo (*Cu*) que é dominante sobre um alelo (*cu*) para orelhas normais. A cor preta resulta de um alelo que se separou de forma independente (*G*) que é dominante sobre um alelo cinza (*g*). Um gato cinza homozigoto com orelhas curvas é cruzado com um gato preto homozigoto com orelhas normais. Todos os gatos de F_1 são pretos e têm orelhas curvas.
   a. Se dois dos gatos de F_1 cruzam, quais fenótipos e proporções são esperados em F_2?
   b. Um gato de F_1 cruza com um gato de rua que é cinza e tem orelhas normais. Quais fenótipos e proporções da prole são esperados a partir deste cruzamento?

*34. Dois genótipos são cruzados: *Aa Bb Cc dd Ee* × *Aa bb Cc Dd Ee*. Qual será a proporção dos seguintes genótipos entre os descendentes deste cruzamento?
   a. *Aa Bb Cc Dd Ee*.
   b. *Aa bb Cc dd ee*.
   c. *aa bb cc dd ee*.
   d. *AA BB CC DD EE*.

35. Nos camundongos, um alelo para olhos cor laranja-amarelado (*a*) é recessivo a um alelo para olhos marrons (a^+). Em um *locus* que se separa de forma independente, um alelo para cor castanha do pelo (*t*) é recessivo para um alelo de cor preta (t^+). Um camundongo homozigoto para olhos marrons e pelo preto é cruzado com um camundongo com olhos laranja-amarelados e pelo castanho. F_1 resultante é intercruzada para produzir F_2. Em uma ninhada de oito camundongos de F_2, qual é a probabilidade de que dois tenham olhos laranja-amarelados e pelo castanho?

36. Nos pepinos, uma fruta fosca (*D*) é dominante sobre a fruta brilhante (*d*), a fruta laranja (*R*) é dominante sobre a fruta creme (*r*) e cotilédones amargos (*B*) são dominantes sobre os cotilédones não amargos (*b*). As três características são codificadas por genes localizados em diferentes pares dos cromossomos. Uma planta homozigota para fruta fosca e laranja e cotilédones amargos é cruzada com uma planta que tem fruta brilhante, de cor creme e cotilédones não amargos. Elas são intercruzadas para produzir F_2.
   a. Indique os fenótipos e suas proporções esperadas em F_2.
   b. Uma planta F_1 é cruzada com uma planta que tem fruta brilhante, de cor creme e cotilédones não amargos. Indique os fenótipos e proporções esperadas entre os descendentes deste cruzamento.

*37. Os alelos *A* e *a* estão localizados em um par de cromossomos metacêntricos. Os alelos *B* e *b* estão localizados em um par de cromossomos acrocêntricos. É feito um cruzamento entre os indivíduos com os seguintes genótipos: *Aa Bb* × *aa bb*.
   a. Esboce os cromossomos como eles apareceriam em cada tipo de gameta produzido pelos indivíduos neste cruzamento.

b. Para cada tipo de descendente resultante deste cruzamento, esboce os cromossomos como eles apareceriam em uma célula em G_1, G_2 e na metáfase da mitose.

Seção 3.4

*38. J. A. Moore investigou a herança dos padrões de mancha nas rãs-leopardo (J. A. Moore. 1943. *Journal of Heredity* 34:3-7). O fenótipo pipiens tem as manchas normais que dão origem ao nome das rãs-leopardo. Ao contrário, o fenótipo burnsi não apresenta manchas no dorso. Moore fez os seguintes cruzamentos, produzindo a prole indicada

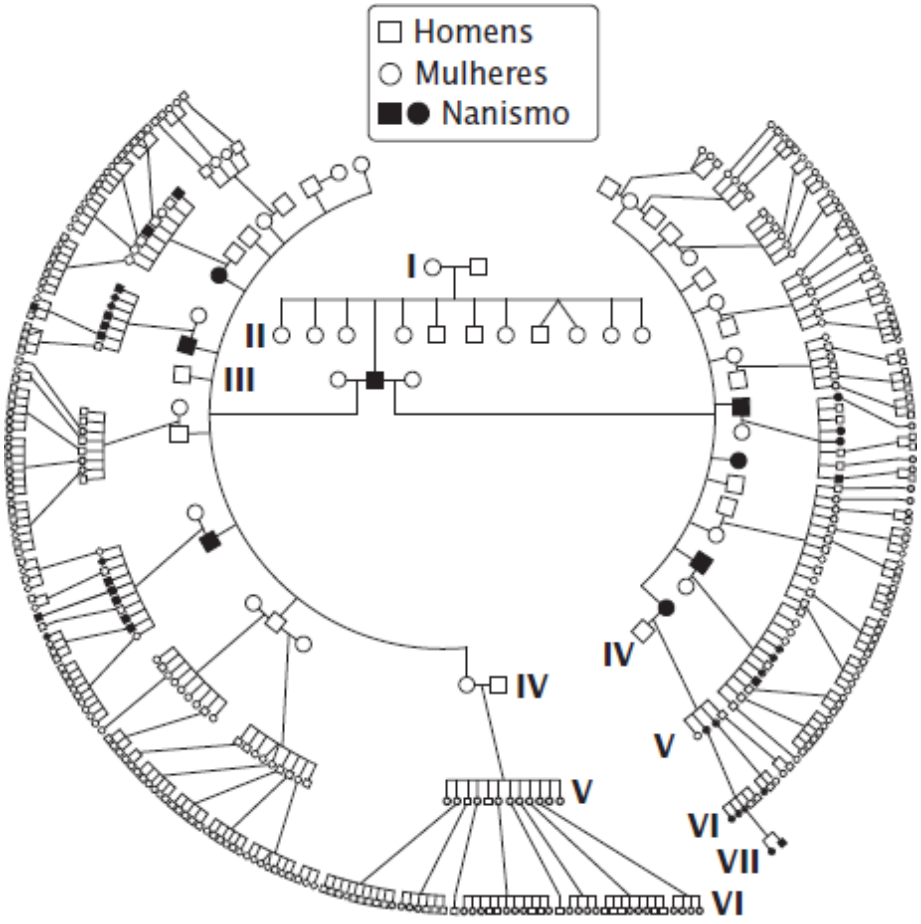
Fenótipos dos genitores	Fenótipos da prole
Burnsi × burnsi	39 burnsi, 6 pipiens
Burnsi × pipiens	23 burnsi, 33 pipiens
Burnsi × pipiens	196 burnsi, 210 pipiens

a. Com base nestes resultados, qual é o modo mais provável de herança do fenótipo burnsi?
b. Indique os genótipos mais prováveis do genitor em cada cruzamento (use *B* para o alelo burnsi e B^+ para o alelo pipiens).
c. Use um teste qui-quadrado para avaliar o ajuste dos números observados da prole para o número esperado com base nos genótipos que você propôs.

39. Nos anos de 1800, um homem com nanismo que viveu em Utah teve vários descendentes: 22 filhos, 49 netos e 250 tataranetos (ver a ilustração de uma linhagem da família à direita), muitos dos quais também eram anões (F. F. Stephens. 1943. *Journal of Heredity* 34:229-235). O tipo de nanismo encontrado nessa família é chamado de condrodisplasia metafisária do tipo Schmid, embora originalmente se acreditasse que fosse nanismo acondroplásico. Entre as famílias destes parentes, o nanismo aparecia apenas em membros que tinham pai ou mãe com nanismo. Quando o pai ou a mãe era anão, produziam-se os seguintes números de crianças.

Família na qual pai ou mãe tinha nanismo	Crianças com estatura normal	Crianças com nanismo
A	15	7
B	4	6
C	1	6
D	6	2
E	2	2
F	8	4
G	4	4
H	2	1
I	0	1
J	3	1
K	2	3
L	2	1
M	2	0
N	1	0
O	0	2
Total	52	40

a. Considerando que a condrodisplasia metafisária do tipo Schmid é rara, esse tipo de nanismo é herdado como um traço dominante ou recessivo? Explique seu raciocínio.
b. Com base na sua resposta para parte *a*, qual é a razão esperada de crianças anãs e normais nas famílias listadas no quadro? Use um teste qui-quadrado para determinar se o número total de crianças destas famílias (52 normais, 40 anãs) é muito diferente do número esperado.



Adaptada de The Journal of Heredity 34:232.

c. Use os testes de qui-quadrado para determinar se o número de crianças na família C (1 normal, 6 anões) e o número na família D (6 normais e 2 anões) são muito diferentes dos números esperados com base no tipo de herança que você propôs. Como você explicaria estes desvios da razão geral esperada?

ANÁLISE DE DADOS

40. O olho cor-de-rosa e albino são dois traços recessivos encontrados no roedor silvestre *Peromyscus maniculatus*. Nos roedores com olhos cor-de-rosa, o olho não tem cor e aparece rosa por causa dos vasos sanguíneos dentro dele. Os roedores albinos não têm cor no seu pelo e olhos. F. H. Clark cruzou os roedores de olhos cor-de-rosa com roedores albinos, a F_1 resultante tinha coloração normal do pelo e olhos. Posteriormente, ele cruzou os roedores de F_1 com roedores com olhos cor-de-rosa e albinos e obteve os seguintes filhotes. É difícil distinguir entre os roedores que são albinos e os que são olhos cor-de-rosa e albinos, então ele combinou estes dois fenótipos (F. H. Clark. 1936. *Journal of Heredity* 27:259-260).

Fenótipo	**Número da prole**
Pelo tipo selvagem, cor dos olhos tipo selvagem	12
Pelo tipo selvagem, olhos cor-de-rosa	62
albino Albino, olhos cor-de-rosa	78
Total	152

a. Indique o número esperado da prole com cada fenótipo se os genes para olhos cor-de-rosa e albinismo se segregam de forma independente.

b. Use um teste qui-quadrado para determinar se o número observado da prole se ajusta ao número esperado com a segregação independente.

*41. Na papoula-da-califórnia, um alelo para flores amarelas (*C*) é dominante sobre um alelo para flores brancas (*c*). Em um *locus* de segregação independente, um alelo para pétalas inteiras (*F*) é dominante sobre um alelo para pétalas com franjas (*f*). Uma planta homozigota para pétalas amarelas e inteiras é cruzada com uma planta com pétalas brancas e com franjas. Uma planta F_1 resultante é então cruzada com uma planta de pétalas brancas e com franjas, produzindo os seguintes

descendentes: 54 com pétalas amarelas e inteiras; 58 com pétalas amarelas e com franjas; 53 com pétalas brancas e inteiras; e 10 com pétalas brancas e com franjas.

a. Use um teste qui-quadrado para comparar os números observados com os números esperados do cruzamento.
b. Que conclusão você pode tirar dos resultados do teste qui-quadrado?
c. Explique seus resultados.

Questões desafiadoras

Seção 3.2

42. O nanismo é um traço recessivo no gado Hereford. Um boiadeiro na região ocidental do Texas descobre que vários bezerros do seu rebanho são anões e ele quer eliminar este traço indesejável do seu rebanho o mais rápido o possível. Suponha que o boiadeiro o contrate como consultor de genética para orientá-lo sobre como retirar o nanismo do seu rebanho. Que cruzamentos você indicaria ao boiadeiro que garantiria que o alelo responsável pelo nanismo seria eliminado do rebanho?

*43. Um geneticista encontrou um camundongo obeso na colônia de seu laboratório; Ele cruzou este camundongo obeso com um camundongo normal. Todos os camundongos de F_1 deste cruzamento tinham peso normal. Quando ele cruzou dois camundongos de F_1, oito camundongos de F_2 tinham peso normal e dois eram obesos. O geneticista intercruzou dois de seus camundongos obesos e descobriu que todos os descendentes deste cruzamento eram obesos. A partir destes resultados, o geneticista concluiu que a obesidade nos camundongos era oriunda de um alelo recessivo.

Uma segunda geneticista em uma diferente universidade também encontrou um camundongo obeso na colônia do seu laboratório. Ela fez os mesmos cruzamentos que o primeiro geneticista fez e obteve os mesmos resultados. Ela também concluiu que a obesidade nos camundongos originava-se de um alelo recessivo. Um dia, os dois geneticistas se encontraram em uma conferência de genética, aprenderam com seus experimentos e decidiram trocar os camundongos. Eles descobriram que, quando cruzavam dois camundongos obesos de diferentes laboratórios, todos os descendentes eram normais; entretanto, quando cruzavam dois camundongos obesos do mesmo laboratório, todos os descendentes eram obesos. Explique estes resultados.

44. O albinismo é um traço recessivo nos seres humanos (ver a introdução do Capítulo 1). Um geneticista estudou várias famílias nas quais os pais eram normais e pelo menos uma das crianças tinham albinismo. O geneticista argumentou que ambos os pais nessas famílias têm de ser heterozigotos e que o albinismo deve aparecer em 1/4 das crianças destas famílias. Para sua surpresa, o geneticista descobriu que a frequência de albinismo nas crianças dessas famílias é muito maior que 1/4. Você consegue explicar a frequência maior que a esperada do albinismo nessas famílias?

45. Dois fenótipos diferentes foram encontrados na salamandra *Plethodon cinereus*: uma forma vermelha e uma preta. Alguns biólogos especularam que o fenótipo vermelho é causado por um alelo autossômico que é dominante sobre um alelo para preto. Infelizmente, essas salamandras não cruzam no cativeiro, então a hipótese de que o vermelho é dominante sobre o preto nunca foi testada.

Um dia, um estudante de genética está caminhando na floresta quando encontra 30 salamandras fêmeas, algumas vermelhas e pretas, colocando ovos. O estudante coloca cada fêmea e seus ovos (cerca de 20 a 30 ovos por fêmea) em bolsas plásticas separadas e leva para o laboratório. Ele consegue incubar os ovos até eclodirem no laboratório. Após os ovos eclodirem, ele registrou os fenótipos das salamandras jovens, junto com os fenótipos de suas genitoras. Assim, o estudante tinha os fenótipos de 30 fêmeas e seus descendentes, mas não tinha informações sobre os fenótipos dos machos.

Explique como o estudante pode determinar se vermelho é dominante sobre preto com estas informações sobre os fenótipos das fêmeas e de seus descendentes.

4

Sexo | Determinação e Características

Curiosidades sobre o sexo do ornitorrinco

O ornitorrinco, *Ornithorhynchus anatinus*, é um dos animais mais estranhos do mundo. Ele é peludo como um mamífero, tem sangue quente e produz leite para seus filhotes, mas não tem dentes, tem bico e bota ovos como um pássaro. Tem membranas digitais nos pés como um pato, e as fêmeas não têm mamilos (os filhotes sugam o leite diretamente da pele do abdome); os machos têm esporões nas patas posteriores que liberam um veneno mortal, como as cobras. O ornitorrinco apresenta uma mistura tal de traços de mamíferos, aves e répteis que os primeiros cientistas a examinarem uma amostra da sua pele pensaram que ele pudesse ser um trote, produzido ao prender partes retiradas de diversos organismos diferentes. Apesar da sua estranha aparência, o ornitorrinco é geneticamente um mamífero monotremado, um ramo que divergiu do resto dos mamíferos há 166 milhões de anos.

Ele vive nas regiões oriental e sul da Austrália e na ilha da Tasmânia. É um excelente nadador, passa boa parte do seu tempo em pequenos rios e riachos procurando vermes, sapos, larvas de insetos, camarões e lagostins. Entre suas peculiaridades, ele localiza sua presa ao detectar correntes elétricas que eles produzem (eletrorrecepção). Seu genoma foi sequenciado em 2008, o que forneceu uma visão detalhada da composição genética deste estranho animal. Ele tem um genoma relativamente pequeno para um mamífero, com 2,3 bilhões de pares de bases de DNA e cerca de 18.500 genes que codificam proteínas. Quase 10% dos seus genes codifica proteínas que são responsáveis pela recepção do odor e substâncias químicas. O genoma do ornitorrinco é uma mistura única de características de mamíferos e répteis.

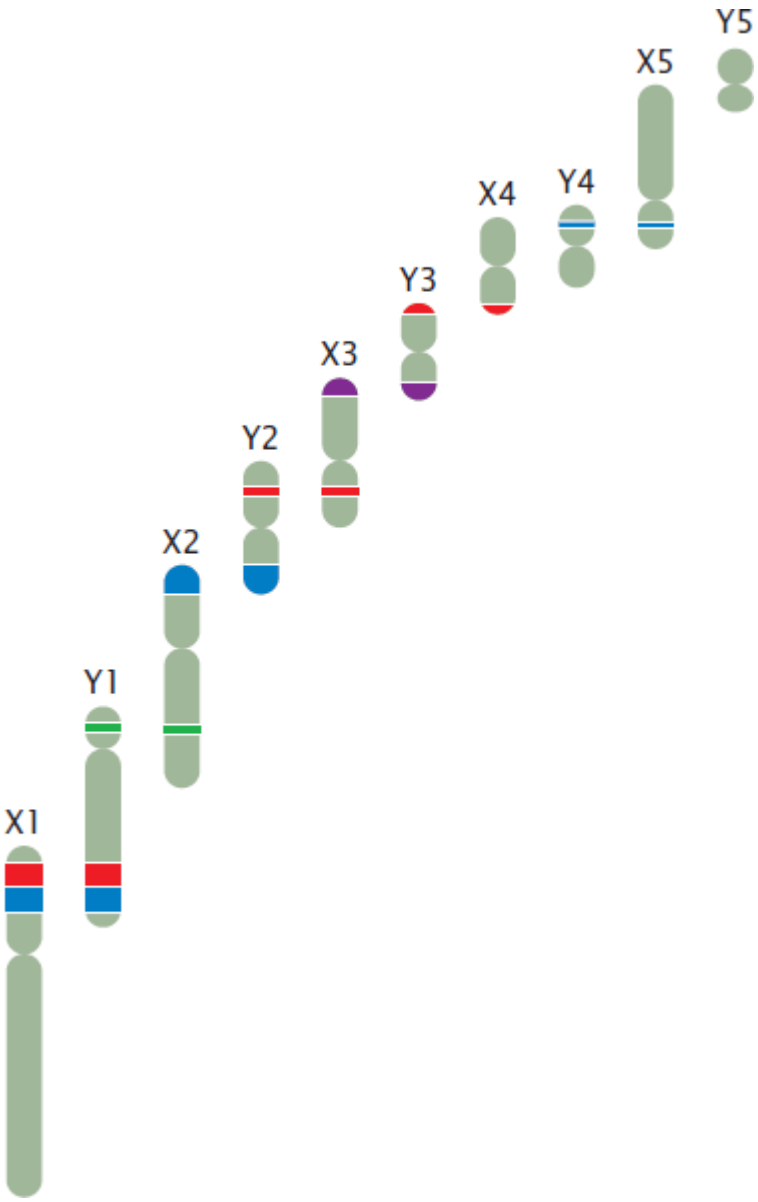


Figura 4.1 Os cromossomos sexuais no ornitorrinco. Na meiose, os cromossomos sexuais formam estruturas semelhantes a cadeias. (Adaptada de F. Veyrunes et al., Genome Research 18(6): 965-973, 2008. Direitos autorais © 2008, Cold Spring Harbor Laboratory Press.)

Seu sexo também é incomum. Para a maior parte dos mamíferos, os cromossomos sexuais determinam se um organismo é masculino ou feminino. As fêmeas têm dois cromossomos X, enquanto os machos têm um único cromossomo X e um cromossomo sexual menor chamado de Y. Este é o tipo comum de determinação de gênero nos mamíferos, mas como o sexo é determinado no ornitorrinco permaneceu um mistério por muitos anos. Este animal tem 52 cromossomos, e os primeiros geneticistas observaram uma mistura de diferentes cromossomos entre os machos e fêmeas desta espécie, incluindo um grupo incomum de cromossomos em cadeia na meiose (**Figura 4.1**).

Em 2004, Frank Grutzner e um grupo de cientistas criaram corantes florescentes para marcar os cromossomos do ornitorrinco de modo que eles pudessem acompanhar o comportamento dos cromossomos durante a meiose. O que eles descobriram foi incrível: os ornitorrincos têm dez cromossomos sexuais; a fêmea tem dez cromossomos X, enquanto o macho tem cinco cromossomos X e cinco cromossomos Y. Na meiose, estes cromossomos sexuais se alinham de forma precisa, formando uma longa cadeia de cromossomos sexuais. Apesar de parecer uma grande confusão, os cromossomos sexuais do ornitorrinco pareiam e se alinham com grande precisão, de modo que cada óvulo tem exatamente cinco Xs, metade dos espermatozoides recebe cinco Xs e a outra metade recebe cinco Ys. O mecanismo que permite esta separação precisa ainda é desconhecido. O conjunto complicado dos cromossomos sexuais no ornitorrinco é apenas um exemplo das formas variadas de determinação do sexo e influencia a hereditariedade.

Estudamos os princípios de segregação e segregação independente de Mendel no Capítulo 3 e vimos como eles explicam a natureza da herança. Após os princípios de Mendel serem redescobertos em 1900, os biólogos começaram a fazer estudos genéticos em uma ampla gama de organismos. Foram observadas exceções à medida que eles aplicavam os princípios de Mendel de forma mais abrangente e se tornou necessário desenvolver extensões dos seus princípios básicos de hereditariedade.

Vamos explorar neste capítulo uma das principais extensões dos princípios de Mendel: a herança de características codificadas por genes localizados nos cromossomos sexuais, que são diferentes nos machos e nas fêmeas. Essas características e os genes que as produzem são chamados de ligados ao sexo. Para entender a herança das características ligadas ao sexo, primeiro precisamos saber como o sexo é determinado – por que alguns membros de uma espécie são machos e outros são fêmeas. A primeira parte deste capítulo é dedicada à determinação do sexo. A segunda parte examina como as características codificadas pelos genes nos cromossomos sexuais são herdadas. Vamos explorar no Capítulo 5 outras formas nas quais o sexo e a herança interagem.

É necessário pensar em dois princípios importantes para estudar a determinação do sexo e as características ligadas a ele. Primeiro, existem vários mecanismos diferentes de determinação do sexo, que podem controlar a herança das características ligadas a ele. Segundo, como os outros pares de cromossomos, os cromossomos sexuais X e Y pareiam no curso da meiose e se separam, mas, na maior parte da sua extensão eles não são homólogos (suas sequências de genes não codificam as mesmas características): a maior parte dos genes no cromossomo X é diferente dos genes no cromossomo Y. Consequentemente, machos e fêmeas não têm os mesmos números de alelos nos *loci* ligados ao sexo. Essa diferença produz padrões variados de herança nos machos e fêmeas. Resolva o Problema 14

4.1 O sexo é determinado por vários mecanismos diferentes

A reprodução sexuada é a formação de descendentes que são geneticamente diferentes de seus genitores; na maior parte dos casos, os genitores contribuem com os genes para seus descendentes e eles são ordenados em novas combinações por meio da meiose. Entre grande parte dos eucariotos, a reprodução sexuada tem dois processos que levam a alternância de células haploides e diploides: a meiose produz gametas haploides (esporos nas plantas) e a fertilização produz zigotos diploides (**Figura 4.2**).

O termo **sexo** se refere ao fenótipo sexual. A maior parte dos organismos tem apenas dois fenótipos sexuais: macho e fêmea. A diferença fundamental entre machos e fêmeas é o tamanho do gameta: os machos produzem gametas menores; as fêmeas produzem gametas relativamente maiores.

O mecanismo pelo qual o sexo é estabelecido é chamado de **determinação do sexo**. Definimos o sexo de um organismo com relação a seu fenótipo. Em alguns casos, um organismo tem cromossomos ou genes que estão normalmente associados a um sexo, mas a anatomia é correspondente ao sexo oposto. Por exemplo, as células das mulheres normalmente têm dois cromossomos X e as células dos homens têm um cromossomo X e um cromossomo Y. Algumas raras pessoas têm anatomia masculina, embora suas células tenham dois cromossomos X. Embora, do ponto de vista genético, essas pessoas sejam mulheres, nós as chamamos de homens porque seu fenótipo sexual é masculino. (Como veremos adiante neste capítulo, estes homens XX em geral têm um pequeno fragmento do cromossomo Y que está preso a outro cromossomo).

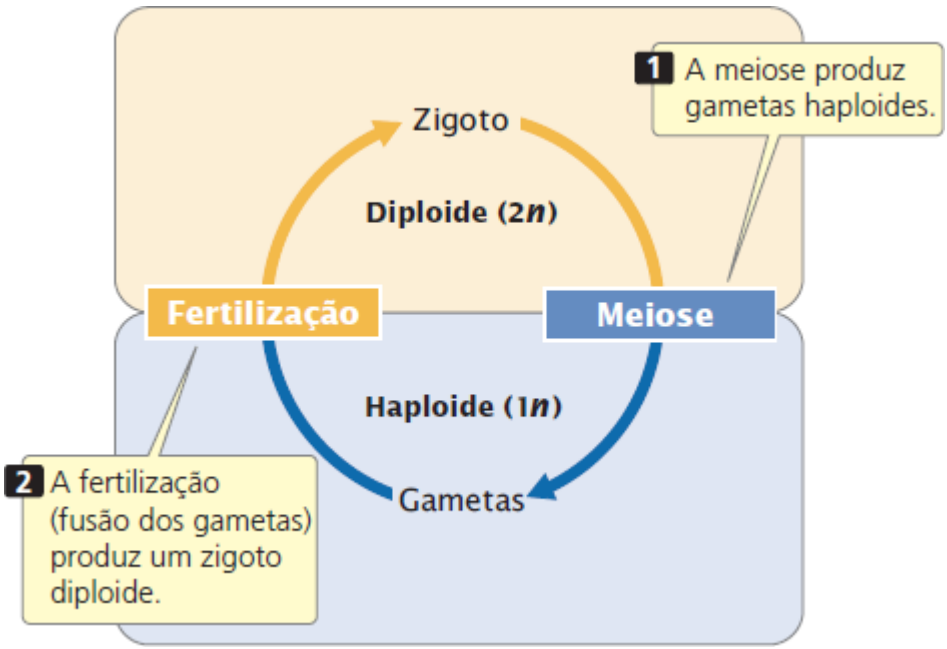


Figura 4.2 Na maior parte dos organismos eucarióticos, a reprodução sexuada é uma alternância de células haploides (1*n*) e diploides (2*n*).

Conceitos

Na reprodução sexuada, os genitores contribuem com genes para produzir descendentes que são geneticamente diferentes deles. Na maior parte dos eucariotos, a reprodução sexuada consiste em meiose, que produz gametas haploides (ou esporos), e fertilização, que produz um zigoto diploide.

Checagem dos conceitos 1

Que processo provoca a variação genética observada nos descendentes produzidos pela reprodução sexuada?

Existem muitas formas nas quais as diferenças sexuais surgem. Em algumas espécies, ambos os sexos são encontrados no mesmo organismo, uma condição chamada de **hermafroditismo**; organismos que têm as estruturas reprodutivas masculinas e femininas são chamados de **monoicos** (hermafroditos) (que significa "uma casa"). As espécies nas quais o organismo tem as estruturas reprodutoras masculina ou feminina são chamadas de **dioicas** ("duas casas"). Os seres humanos são dioicos. Nas espécies dioicas, o sexo pode ser determinado pelo cromossomo, pela genética ou pelo meio ambiente.

Sistemas cromossômicos para determinação do sexo

A teoria cromossômica da hereditariedade (ver *Capítulo 3*) afirma que os genes estão localizados nos cromossomos, que servem como veículo para a sua separação na meiose. A prova definitiva desta teoria foi a descoberta de que o sexo de alguns insetos é determinado pela presença ou ausência de cromossomos específicos.

Em 1891, Hermann Henking observou uma estrutura peculiar nos núcleos das células dos insetos machos. Embora não compreendesse sua função nem sua relação com o sexo, ele chamou esta estrutura de corpo X. Posteriormente, Clarence E. McClung estudou o corpo X nos gafanhotos e identificou que se tratava de um cromossomo. McClung chamou-o de cromossomo acessório, mas ele ficou conhecido como cromossomo X, por causa do nome original dado por Henking. McClung observou que as células dos gafanhotos-fêmeas tinham um cromossomo a mais que o número de cromossomos nas células dos machos e concluiu que os cromossomos acessórios participavam na determinação do sexo. Em 1905, Nettie Stevens e Edmund Wilson demonstraram que, nos gafanhotos e em outros insetos, as células das fêmeas tinham dois cromossomos X, enquanto as células dos machos tinham apenas um único X. Em alguns insetos, eles contaram o mesmo número de cromossomos nas células dos machos e das fêmeas, mas viram que um par de cromossomos era diferente: encontravam dois cromossomos X nas células das fêmeas e um único cromossomo X mais um cromossomo menor, que eles chamaram de Y, nas células dos machos.

Stevens e Wilson também mostraram que os cromossomos X e Y se separavam para diferentes células na formação dos espermatozoides, metade dos espermatozoides recebia um cromossomo X e a outra metade recebia um Y. Todos os óvulos produzidos pelas fêmeas na meiose recebiam um cromossomo X. Um espermatozoide com um cromossomo Y se une a um óvulo que carreia um cromossomo X para produzir um macho XY, enquanto um espermatozoide com cromossomo X se une a um óvulo com cromossomo X para produzir uma fêmea XX. Esta distribuição dos cromossomos X e Y no espermatozoide é responsável pela razão 1:1 do sexo observada nos organismos dioicos (**Figura 4.3**). Da mesma forma que o sexo é herdado como outras características geneticamente determinadas, a descoberta de Stevens e Wilson de que o sexo está associado à herança de um cromossomo específico também demonstrou que os genes estão nos cromossomos.

Assim como Stevens e Wilson encontraram nos insetos, o sexo em muitos organismos é determinado por um par de cromossomos, os **cromossomos sexuais**, que diferem entre machos e fêmeas. Os cromossomos não sexuais, que são os mesmos para machos e fêmeas, são chamados de **autossomos**. Pensamos no sexo nos organismos com cromossomos sexuais sendo determinado pela presença destes cromossomos, mas, na verdade, os genes individuais localizados neles são responsáveis pelos fenótipos sexuais.

Determinação do sexo pelo sistema XX-XO. O mecanismo para determinação do sexo nos gafanhotos estudado por McClung é um dos mais simples para a determinação do cromossomo sexual e é chamado de sistema XX-XO. Neste sistema, as fêmeas têm dois cromossomos X (XX), e os machos têm um único cromossomo X (XO). Não existe o cromossomo O – a letra O indica a ausência de um cromossomo sexual.

Durante a meiose nas fêmeas, os dois cromossomos X se pareiam e se separam, com um cromossomo indo para cada óvulo haploide. Nos machos, o único cromossomo X se separa na meiose para metade dos espermatozoides, a outra metade não recebe o cromossomo sexual. Como os machos produzem dois diferentes tipos de gametas com os cromossomos sexuais, eles são considerados o **sexo heterogamético**. As fêmeas que produzem gametas com os mesmos cromossomos sexuais são o **sexo homogamético.** No sistema XX-XO, o sexo de um organismo é, portanto, determinado por qual tipo de gameta masculino que fertiliza o óvulo. Os espermatozoides com X se unem aos óvulos com X para produzir zigotos XX, que acabam se tornando como fêmeas. Os espermatozoides que não têm o cromossomo X se unem aos óvulos com X para produzir zigotos XO, que se

desenvolvem em machos.

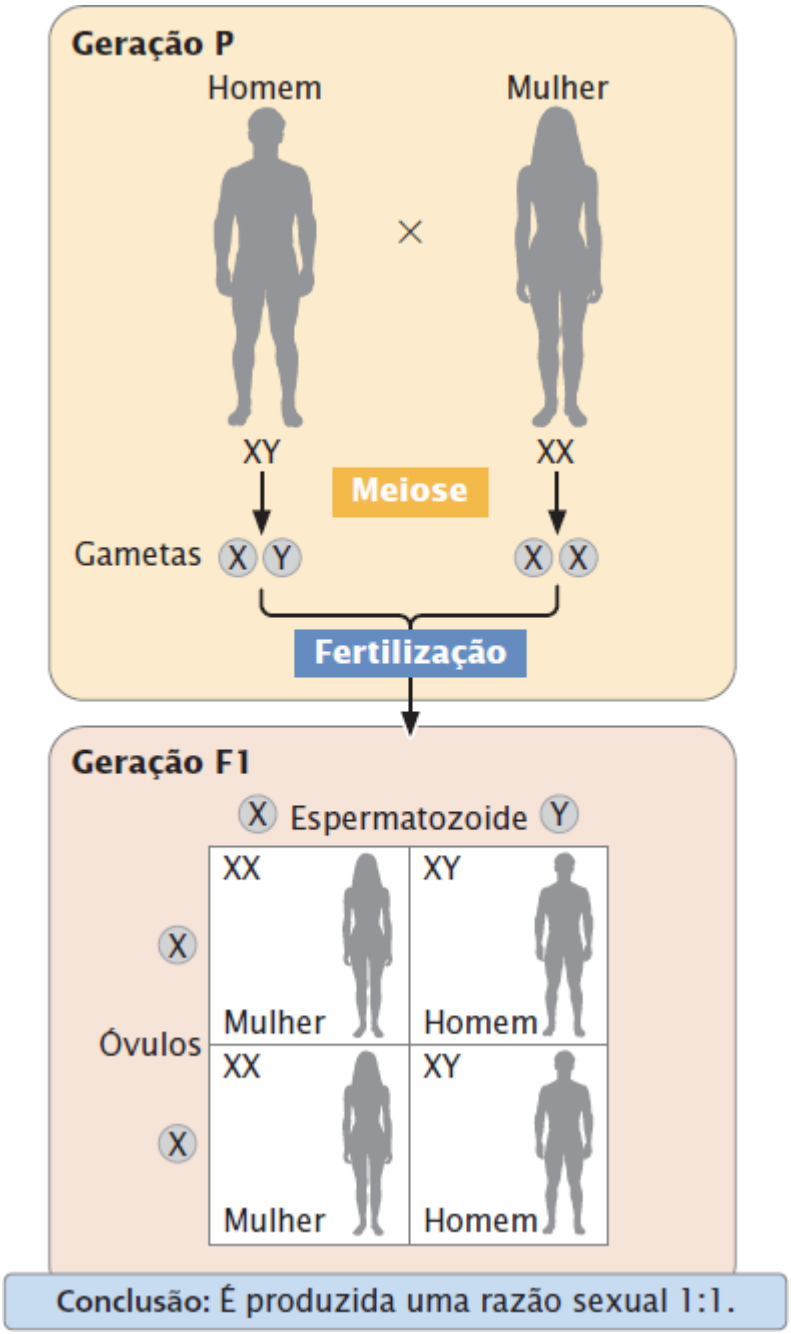


Figura 4.3 A herança do sexo nos organismos com cromossomos X e Y resulta em um número igual de descendentes masculinos e femininos.

Determinação do sexo pelo sistema XX-XY. Em muitas espécies, as células dos machos e fêmeas têm o mesmo número de cromossomos, mas as células das fêmeas têm dois cromossomos X (XX) e as células dos machos têm um único cromossomo X e um cromossomo sexual menor, o cromossomo Y (XY). Nos seres humanos e em muitos outros organismos, o cromossomo Y é acrocêntrico (**Figura 4.4**), e não no formato Y como frequentemente é presumido. Neste tipo de sistema de determinação de sexo, o macho é o sexo heterogamético – metade dos seus gametas tem um cromossomo X e a outra metade, um cromossomo Y. A fêmea é o sexo homogamético – todos seus óvulos têm um único cromossomo X. Muitos organismos, incluindo algumas plantas, insetos, répteis e mamíferos (entre eles os seres humanos) têm o sistema XX-XY de determinação do sexo. Outros organismos têm variações do sistema XX-XY, como os ornitorrincos (discutidos na introdução deste capítulo) nos quais as fêmeas têm cinco pares de cromossomos X e os machos, cinco pares de cromossomos X e Y.

Embora os cromossomos X e Y não sejam homólogos, eles se pareiam e se separam em diferentes células na meiose. Eles conseguem se parear porque seus cromossomos são homólogos em pequenas regiões chamadas de **regiões pseudoautossômicas** (ver Figura 4.4), onde eles carreiam os mesmos genes. Existem regiões pseudoautossômicas em ambas as extremidades dos cromossomos X e Y nos seres humanos.

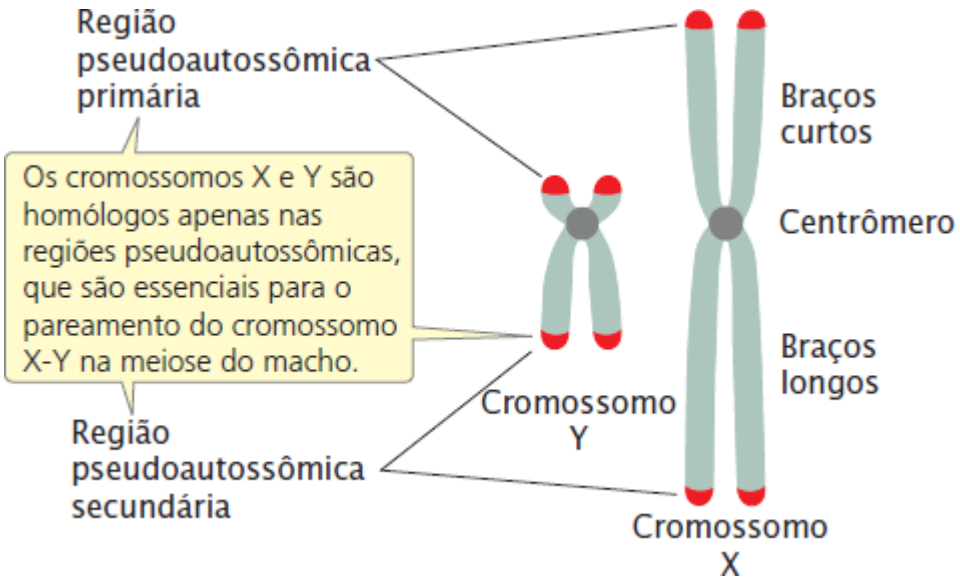


Figura 4.4 O tamanho e o conteúdo dos cromossomos X e Y nos seres humanos são diferentes. Eles são homólogos apenas nas regiões pseudoautossômicas.

Determinação do sexo pelo sistema ZZ-ZW. Neste sistema, a fêmea é heterogamética e o macho é homogamético. Para evitar confusão com o sistema XX-XY, os cromossomos sexuais neste sistema são chamados de Z e W, mas os cromossomos não lembram tais letras. As fêmeas neste sistema são ZW; após a meiose, metade dos óvulos tem um cromossomo Z e a outra metade tem um cromossomo W. Os machos são ZZ, todos os espermatozoides têm um único cromossomo Z. O sistema ZZ-ZW é encontrado nas aves, cobras, borboletas, alguns anfíbios e alguns peixes.

Conceitos

A descoberta de que a presença ou ausência de cromossomos específicos determina o sexo nos insetos forneceu evidência de que os genes estão localizados nos cromossomos. Na determinação do sexo pelo sistema XX-XO, o macho é XO e heterogamético e a fêmea é XX e homogamética. Na determinação do sexo pelo XX-XY, o macho é XY e a fêmea é XX; o macho é heterogamético neste sistema. Na determinação do sexo pelo ZZ-ZW, a fêmea é ZW e o macho é ZZ; a fêmea é heterogamética neste sistema.

Checagem dos conceitos 2

Qual a diferença entre sexo heterogamético e sexo homogamético?

a. O sexo heterogamético é o macho; o sexo homogamético é a fêmea.
b. Os gametas do sexo heterogamético têm diferentes cromossomos sexuais; os gametas do sexo homogamético têm o mesmo cromossomo sexual.
c. Todos os gametas do sexo heterogamético têm um cromossomo Y; todos os gametas do sexo homogamético têm um cromossomo X.

Determinação do sexo pelo gene

Em alguns organismos, o sexo é determinado geneticamente, mas não existem diferenças claras nos cromossomos dos machos e fêmeas; ou seja, não existem cromossomos sexuais. Estes organismos têm **determinação do sexo pelo gene**, os genótipos em um ou mais *loci* determinam o sexo de um indivíduo. Os cientistas observaram a determinação do sexo pelo gene em algumas plantas, fungos, protozoários e peixes.

É importante compreender que, mesmo nos sistemas cromossômicos de determinação de sexo, este, de fato é determinado por genes individuais. Por exemplo, nos mamíferos, um gene (*SRY*, discutido adiante neste capítulo) localizado no cromossomo Y, determina o fenótipo masculino. Na determinação do sexo pelo gene e pelos cromossomos, o sexo é controlado por genes individuais, a diferença é que com a determinação do sexo pelos cromossomos, os cromossomos sexuais também são diferentes nos machos e fêmeas.

Determinação do sexo pelo ambiente

Os genes participam em todos os exemplos de determinação do sexo discutidos até o momento, mas, em vários organismos, o

sexo é determinado em parte ou totalmente por fatores ambientais.

Um exemplo fascinante da determinação do sexo pelo ambiente é observado no molusco marinho *Crepidula fornicata*, também conhecida como lapa **(Figura 4.5)**. Estes moluscos vivem em pilhas, um em cima do outro. Cada molusco inicia sua vida como uma larva nadadora. A primeira larva que se deposita em um substrato sólido, não ocupado, torna-se em um molusco-fêmea. Ele então produz substâncias químicas que atraem outras larvas, que se depositam no topo. Estas larvas se tornam machos, que então cruzam com o molusco que está embaixo. Após um período, os machos no topo se tornam fêmeas e, por sua vez, atraem mais larvas que se depositam no topo da pilha, se desenvolvem em machos e cruzam com os moluscos abaixo. Os moluscos podem formar pilhas de dúzias ou mais; os que ficam nas posições mais altas da pilha são sempre os machos. Este tipo de desenvolvimento sexual é chamado de **hermafroditismo sequencial**, em que cada animal pode ser tanto macho quanto fêmea, embora não ao mesmo tempo. No caso do *Crepidula fornicata*, o sexo é determinado pelo meio ambiente por meio da posição do molusco na pilha.

Os fatores ambientais também são importantes para determinar o sexo em alguns répteis; o fenótipo sexual de muitas tartarugas, crocodilos, jacarés e alguns pássaros é afetado pela temperatura durante o desenvolvimento embriogênico. Nas tartarugas, por exemplo, as temperaturas quentes na incubação produzem mais fêmeas, enquanto as temperaturas mais frias produzem machos. O inverso acontece com os jacarés. Em algumas espécies, os cromossomos sexuais geralmente determinam se os indivíduos são machos ou fêmeas, entretanto, alguns fatores ambientais podem se sobrepor à determinação do sexo pelo cromossomo. Por exemplo, os lagartos dragões barbudos machos são normalmente ZZ e as fêmeas são ZW, mas, quando os ovos são incubados em altas temperaturas, os indivíduos ZZ se desenvolvem com fenótipo de fêmea. Alguns dos diferentes tipos de determinação do sexo estão resumidos no **Quadro 4.1**.

Agora que revisamos algumas das diferentes formas pelas quais o sexo pode ser determinado, vamos examinar detalhadamente um mecanismo, o sistema XX-XY. A determinação do sexo é XX-XY nas moscas-da-fruta e seres humanos, mas, como veremos, a forma pela qual os cromossomos X e Y determinam o sexo nesses dois organismos é bem diferente.

Resolva os Problemas 4 e 21

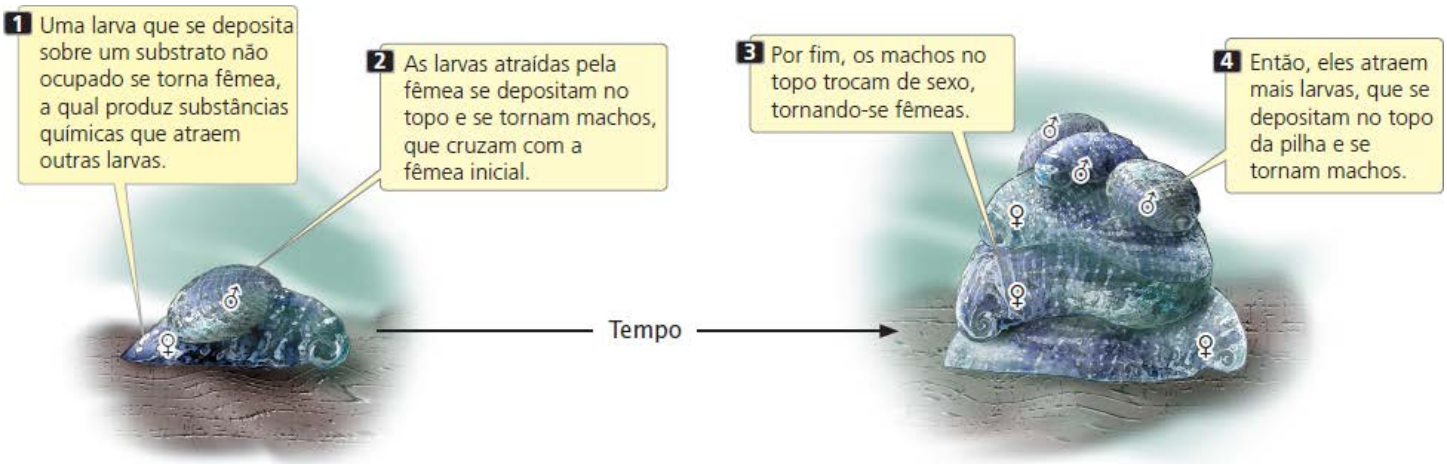


Figura 4.5 No *Crepidula fornicata*, o molusco marinho comumente conhecido como lapa, o sexo é determinado por um fator ambiental – a posição do molusco na pilha.

Quadro 4.1 Sistemas para determinar o sexo.

Sistema	Mecanismo		Sexo heterogamético	Organismos
XX-XO	Fêmeas XX	Machos X	Macho	Alguns gafanhotos e outros insetos
XX-XY	Fêmeas XX	Machos XY	Macho	Muitos insetos, peixes, anfíbios, répteis, mamíferos, incluindo os seres humanos
ZZ-ZW	Fêmeas ZW	Machos ZZ	Fêmea	Borboletas,

				pássaros, alguns répteis e anfíbios
Determinação do sexo pelo gene	Sem cromossomos sexuais diferentes	Sexo determinado pelos genes em cromossomos não diferenciados	Varia	Algumas plantas, fungos, protozoários e peixes
Determinação do sexo pelo ambiente	Sexo determinado pelos fatores ambientais		Nenhum	Alguns invertebrados, tartarugas, jacarés

Conceitos

Na determinação do sexo pelos genes, ele é determinado pelos genes em um ou mais *loci*, mas não existem diferenças claras nos cromossomos dos machos e das fêmeas. Na determinação do sexo pelo ambiente, o sexo é determinado em parte ou totalmente pelos fatores ambientais.

Checagem dos conceitos 3

Qual é a diferença na determinação do sexo pelo cromossomo, gene e meio ambiente?

Determinação do sexo na *Drosophila melanogaster*

A mosca-de-fruta *Drosophila melanogaster* tem oito cromossomos: três pares de autossomos e um par de cromossomos sexuais. Em geral, as fêmeas têm dois cromossomos sexuais X e os machos, um X e um Y. Nos anos 1920, Calvin Bridges propôs que o sexo na *Drosophila* era determinado não pelo número de cromossomos X e Y, mas pelo equilíbrio dos genes que determinam a fêmea no cromossomo X e os genes que determinam o macho nos autossomos. Ele sugeriu que o sexo da mosca é determinado pela razão X:A, o número de cromossomos X dividido pelo número de conjuntos haploides de cromossomos autossômicos. As moscas normais têm dois conjuntos haploides de autossomos e dois cromossomos X (fêmeas) ou um cromossomo X e um cromossomo Y (machos). Bridges propôs que uma razão X:A de 1,0 produz uma mosca-fêmea; uma razão X:A de 0,5 produz uma mosca-macho. Ele também sugeriu que uma razão X:A entre 1,0 e 0,5 produz uma mosca intersexuada, com uma mistura de características de macho e fêmea. Uma razão X:A menor que 0,05 e maior que 1,0 produz moscas com desenvolvimento anormais chamadas de metamachos e metafêmeas, respectivamente. Quando Bridges e outros pesquisadores examinaram as moscas com diferentes números de cromossomos sexuais e autossomos, a razão X:A parecia prever corretamente o sexo fenotípico das moscas (**Quadro 4.2**).

Embora a razão X:A *preveja* corretamente o fenótipo sexual, pesquisa recente sugere que o *mecanismo* de determinação do sexo não é um equilíbrio entre os genes ligados ao X e os autossômicos, como Bridges propôs. Os pesquisadores localizaram vários genes no cromossomo X que afetam o fenótipo sexual, mas foram identificados alguns genes autossômicos que determinam o sexo (necessários para a hipótese de razão X:A). O novo indício sugere que os genes no cromossomo X sejam os determinantes primários do sexo. A influência do número de cromossomos autossômicos no sexo é indireta, afetando a cronologia dos eventos de desenvolvimento, e, portanto, por quanto tempo os genes que determinam o sexo no cromossomo X estão ativos. Por exemplo, as moscas XX com três conjuntos autossômicos (XX, AAA) têm uma razão X:A de 0,67 e desenvolvem um fenótipo intersexuado. A existência de três conjuntos autossômicos leva a um estágio crítico do desenvolvimento, não permitindo que os fatores femininos codificados nos cromossomos X tenham tempo suficiente para se acumularem, fazendo com que estas moscas tenham um fenótipo intersexuado. O número de cromossomos autossômicos influencia a determinação do sexo na *Drosophila*, mas não por meio da ação dos genes autossômicos, como previsto por Bridges.

Quadro 4.2 Complementos de cromossomo e fenótipos sexuais em *Drosophila*.

Complemento do cromossomo sexual	Conjuntos haploides de autossomos	Razão X: A	Fenótipo sexual
XX	AA	1,0	Fêmea

XY	AA	0,5	Macho
XO	AA	0,5	Macho
XXY	AA	1,0	Fêmea
XXX	AA	1,5	Metafêmea
XXXY	AA	1,5	Metafêmea
XX	AAA	0,67	Intersexuado
XO	AAA	0,33	Metamacho
XXXX	AAA	1,3	Metafêmea

Conceitos

Embora o fenótipo sexual de uma mosca-da-fruta seja previsto pela razão X:A, na verdade, o sexo é determinado pelos genes no cromossomo X.

Determinação do sexo nos seres humanos

Os seres humanos, como a *Drosophila*, têm a determinação do sexo pelo XX-XY, mas a presença de um gene (*SRY*) no cromossomo Y determina as características masculinas. Os fenótipos resultantes de vários cromossomos sexuais anormais, que surgem quando estes cromossomos não se separam de forma adequada na meiose ou mitose, ilustram a importância do cromossomo Y na determinação do sexo nos seres humanos.

Síndrome de Turner. As mulheres são o sexo afetado pela **síndrome de Turner** e em muitos casos têm as características sexuais secundárias subdesenvolvidas. Esta síndrome é observada em 1 em cada 3.000 recém-nascidas. As mulheres afetadas têm, em muitos casos, baixa estatura e a linha de implantação do cabelo baixa, um tórax relativamente largo e dobras de pele no pescoço. Em geral, apresentam inteligência normal. A maior parte das mulheres portadoras da síndrome de Turner é estéril. Em 1959, Charles Ford usou novas técnicas para estudar os cromossomos humanos e descobriu que as células de uma menina de 14 anos portadora da síndrome de Turner tinham apenas um único cromossomo X (**Figura 4.6**); este complemento de cromossomo é chamado de XO.

Não existem casos conhecidos de ausência dos dois cromossomos X, uma indicação de que pelo menos um cromossomo X é necessário para o desenvolvimento humano. Aparentemente, os embriões sem os dois X são abortados espontaneamente nos estágios iniciais do desenvolvimento.

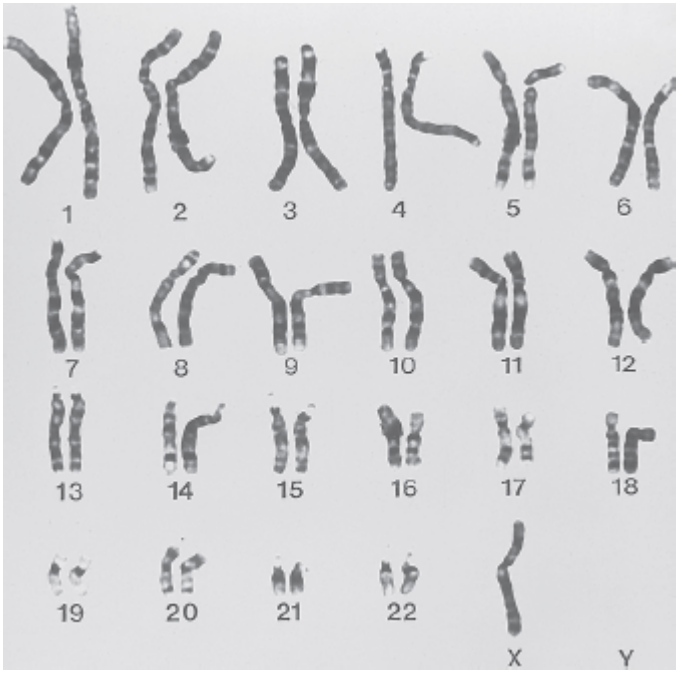


Figura 4.6 Pessoas com síndrome de Turner têm um único cromossomo X nas suas células. (Department of Clinical Cytogenetics, Addenbrookes Hospital/Science Photo Library/Photo Researchers.)

Síndrome de Klinefelter. Os portadores da **síndrome de Klinefelter**, que surge em uma frequência de em torno de 1 em cada 1.000 meninos, têm células com um ou mais cromossomos Y e múltiplos cromossomos X. A maior parte das células dos homens com esta síndrome é XXY (**Figura 4.7**), mas as células também podem ser XXXY, XXXXY ou XXYY. Os homens com essa condição têm testículos pequenos e pelos faciais e púbicos reduzidos. São com frequência mais altos que o normal, estéreis, e a maioria tem inteligência normal.

Mulheres poli-X. Em cerca de 1 em cada 1.000 meninas, as células têm três cromossomos X, um quadro chamado de **síndrome do X triplo**. Estas pessoas não têm características especiais além da tendência a serem altas e magras. Embora algumas sejam estéreis, muitas menstruam regularmente e são férteis. A incidência de déficit intelectual nas mulheres X triplo é ligeiramente maior que na população geral, mas a maior pate das mulheres XXX apresenta inteligência normal. É muito mais raro encontrar mulheres cujas células tenham quatro ou cinco cromossomos X. Em geral elas têm anatomia feminina normal, mas têm comprometimento intelectual e apresentam vários problemas físicos. A gravidade do comprometimento intelectual aumenta à medida que o número de cromossomos excede o número três.

Figura 4.7 As pessoas com síndrome de Klinefelter têm um cromossomo Y e dois ou mais cromossomos X em suas células. (Biophoto Associates/Science Source/Photo Researchers.)

O papel dos cromossomos sexuais. Os fenótipos associados às anomalias dos cromossomos sexuais permitem várias inferências sobre a participação dos cromossomos sexuais na determinação do sexo nos seres humanos.

1. O cromossomo X tem informações genéticas essenciais para ambos os sexos; pelo menos uma cópia de um cromossomo X é necessária para o desenvolvimento humano.
2. O gene que determina o sexo masculino está localizado no cromossomo Y. Uma única cópia deste cromossomo, mesmo na presença de vários cromossomos X, produz fenótipo masculino.
3. A ausência do cromossomo Y geralmente resulta em fenótipo feminino.
4. Os genes que afetam a fertilidade estão localizados nos cromossomos X e Y. Em geral, uma mulher precisa de pelo menos duas cópias do cromossomo X para ser fértil.
5. Cópias adicionais do cromossomo X prejudicam o desenvolvimento normal em homens e mulheres, produzindo problemas físicos e mentais que aumentam à medida que o número de cromossomos X extras aumenta.

O gene para determinação do sexo masculino nos seres humanos. O cromossomo Y nos seres humanos e em todos os outros mamíferos é muito importante para produzir o fenótipo masculino. Entretanto, os cientistas descobriram alguns raros homens XX cujas células aparentemente não têm cromossomo Y. Por muitos anos, estes homens eram um enigma: Como um fenótipo masculino poderia existir sem um cromossomo Y? Um exame detalhado revelou uma pequena parte do cromossomo Y preso a outro cromossomo. Este achado indica que não é o cromossomo Y inteiro que determina a masculinidade nos seres

humanos, mas sim um gene presente nele.

No início do desenvolvimento, todos os seres humanos têm gônadas não diferenciadas e ductos reprodutivos masculinos e femininos. Aproximadamente 6 semanas após a fertilização, um gene no cromossomo Y é ativado. Ele permite que as gônadas neutras se tornem testículos, que começam a secretar dois hormônios: a testosterona e a substância inibidora mülleriana. A testosterona induz ao desenvolvimento das características masculinas e a substância inibidora mülleriana leva à degeneração dos ductos reprodutores femininos. Na ausência desse gene que determina o sexo masculino, as gônadas neutras se tornam ovários e as características femininas se desenvolvem.

O gene que determina o sexo masculino, chamado de **gene da região determinante do sexo no cromossomo Y** (*SRY*), foi descoberto em 1990 (**Figura 4.8**). Esse gene foi encontrado em homens XX e não existe nas mulheres XY, sendo também encontrado no cromossomo Y de outros mamíferos. A prova definitiva de que o *SRY* é o gene determinante do sexo masculino veio quando os cientistas inseriram uma cópia do gene em camundongos XX por meio da engenharia genética. Os camundongos XX que receberam este gene, embora fossem estéreis, desenvolveram características anatômicas dos machos.

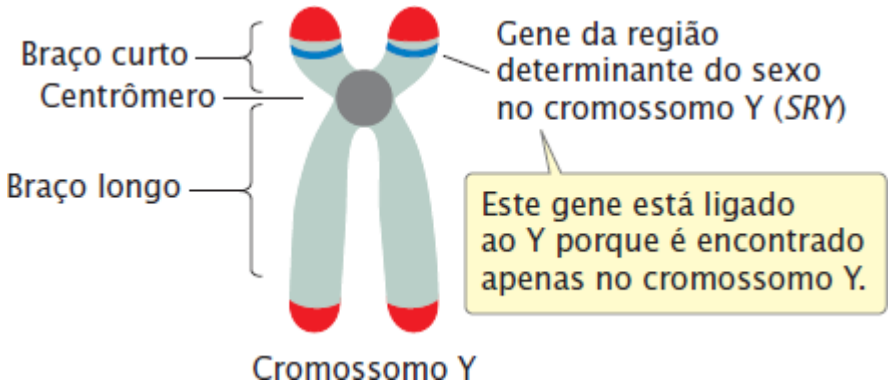


Figura 4.8 O gene *SRY* está no cromossomo Y e leva ao desenvolvimento das características masculinas.

O gene *SRY* codifica uma proteína chamada de fator de transcrição (ver *Capítulo 13*), que se liga ao DNA e estimula a transcrição de outros genes que promovem a diferenciação dos testículos. Embora o *SRY* seja o determinante primário da masculinidade nos seres humanos, outros genes (alguns ligados ao X, outros ligados ao Y e ainda alguns autossomos) também têm papéis na fertilidade e no desenvolvimento das diferenças sexuais.

Conceitos

O gene *SRY* no cromossomo Y faz com que o embrião humano se desenvolva como um homem. Se não existir este gene, o embrião se desenvolve como mulher.

Checagem dos conceitos 4

Qual é o fenótipo de uma pessoa que tem cromossomos sexuais XXXY?

a. Síndrome de Klinefelter.
b. Síndrome de Turner.
c. Mulher poli-X.

Síndrome da insensibilidade aos andrógenos. Embora o gene *SRY* seja o determinante primário do sexo nos embriões humanos, vários outros genes influenciam o desenvolvimento sexual, como ilustrado por mulheres com a síndrome da insensibilidade aos andrógenos. Estas pessoas têm características sexuais externas femininas. De fato, a maior parte não tem consciência da sua condição até alcançarem a puberdade e não menstruar. O exame ginecológico revela que a vagina tem fundo cego e que não há útero, tubas uterinas nem ovários. Na cavidade abdominal, um par de testículos produz níveis de testosterona observados normalmente nos homens. As células de uma mulher com a síndrome da insensibilidade aos andrógenos têm um cromossomo X e um Y.

Como uma pessoa pode ser mulher na aparência quando suas células têm um cromossomo Y e ela tem testículos que produzem testosterona? A resposta está na complexa relação entre os genes e o sexo nos seres humanos. Em um embrião humano com um cromossomo Y, o gene *SRY* faz com que as gônadas se desenvolvam em testículos, que produzem testosterona. A testosterona estimula os tecidos embrionários a desenvolverem as características masculinas. Mas, para que a testosterona exerça seu efeito, ela precisa se ligar a um receptor de androgênio. Este receptor é defeituoso nas mulheres com a síndrome de insensibilidade aos andrógenos; consequentemente, suas células são insensíveis à testosterona e as características

femininas se desenvolvem. O gene para o receptor para andrógenos está localizado no cromossomo X, então estas pessoas sempre herdam esta síndrome de suas mães.

A síndrome da insensibilidade aos andrógenos ilustra questões sobre a influência dos genes no sexo de uma pessoa. Primeiro, esta condição demonstra que o desenvolvimento sexual humano é um processo complexo, influenciado não apenas pelo gene *SRY* no cromossomo Y, mas também por genes localizados em outros cromossomos. Segundo, ela mostra que a maior parte das pessoas tem genes para as características masculinas e femininas, como ilustrado pelo fato de que as pessoas com a síndrome da insensibilidade aos andrógenos desenvolvem características femininas, mesmo tendo cromossomos masculinos. Na verdade, os genes para muitas das características sexuais femininas e masculinas secundárias não são encontrados nos cromossomos sexuais, mas nos autossomos. A chave para a masculinidade e feminilidade não está nos genes, mas sim no controle da sua expressão. **Resolva o Problema 17**

4.2 As características ligadas ao sexo são determinadas por genes nos cromossomos sexuais

Aprendemos no Capítulo 3 alguns princípios básicos da hereditariedade que Mendel descobriu por meio de seus cruzamentos entre as ervilhas. Uma importante extensão desses princípios mendelianos é o padrão de herança exibido pelas **características ligadas ao sexo**, determinadas pelos genes localizados nos cromossomos sexuais. Os genes no cromossomo X determinam as **características ligadas ao X**; os genes no cromossomo Y determinam as **características ligadas ao Y**. Como o cromossomo Y de muitos organismos tem pouca informação genética, a maior parte das características ligadas ao sexo está de fato conectada ao cromossomo X. Homens e mulheres são diferentes nos seus cromossomos sexuais, então o padrão de herança para as características ligadas ao sexo é diferente do exibido pelos genes localizados nos cromossomos autossômicos.

Os olhos brancos ligados ao X na *Drosophila*

A primeira pessoa a explicar a herança ligada ao sexo foi o biólogo americano Thomas Hunt Morgan (**Figura 4.9**). Ele começou sua carreira como embriologista, mas a descoberta dos princípios de Mendel o inspirou a fazer experimentos genéticos, inicialmente em camundongos e ratos. Em 1909, Morgan trocou sua pesquisa para a *Drosophila melanogaster*; 1 ano depois, descobriu entre as moscas da colônia do seu laboratório um único macho que tinha olhos brancos, em total contraste com os olhos vermelhos nas moscas-da-fruta normais. Esta mosca teve um efeito incrível na carreira de Morgan como biólogo e no futuro da genética.

Para investigar a herança da característica de olhos brancos nestas moscas, Morgan realizou de forma sistemática vários cruzamentos genéticos. Primeiro, ele cruzou fêmeas de linhagem pura e olhos vermelhos com macho de olhos brancos, produzindo descendentes F_1 que tinham olhos vermelhos (**Figura 4.10 A**). Os resultados de Morgan para este cruzamento inicial eram consistentes com os princípios de Mendel: um cruzamento entre um indivíduo homozigoto dominante e um indivíduo homozigoto recessivo produz descendentes heterozigotos que exibem o traço dominante. Seus resultados sugeriram que os olhos brancos eram um único traço recessivo. Entretanto, quando Morgan cruzou as moscas de F_1 entre si, ele descobriu que todas as moscas-fêmeas de F_2 tinham olhos vermelhos, mas metade das moscas-machos de F_2 tinha olhos vermelhos e a outra metade tinha olhos brancos. Essa descoberta não era o resultado esperado para um único traço recessivo, que deve obviamente aparecer em 1/4 dos machos e fêmeas dos descendentes F_2.

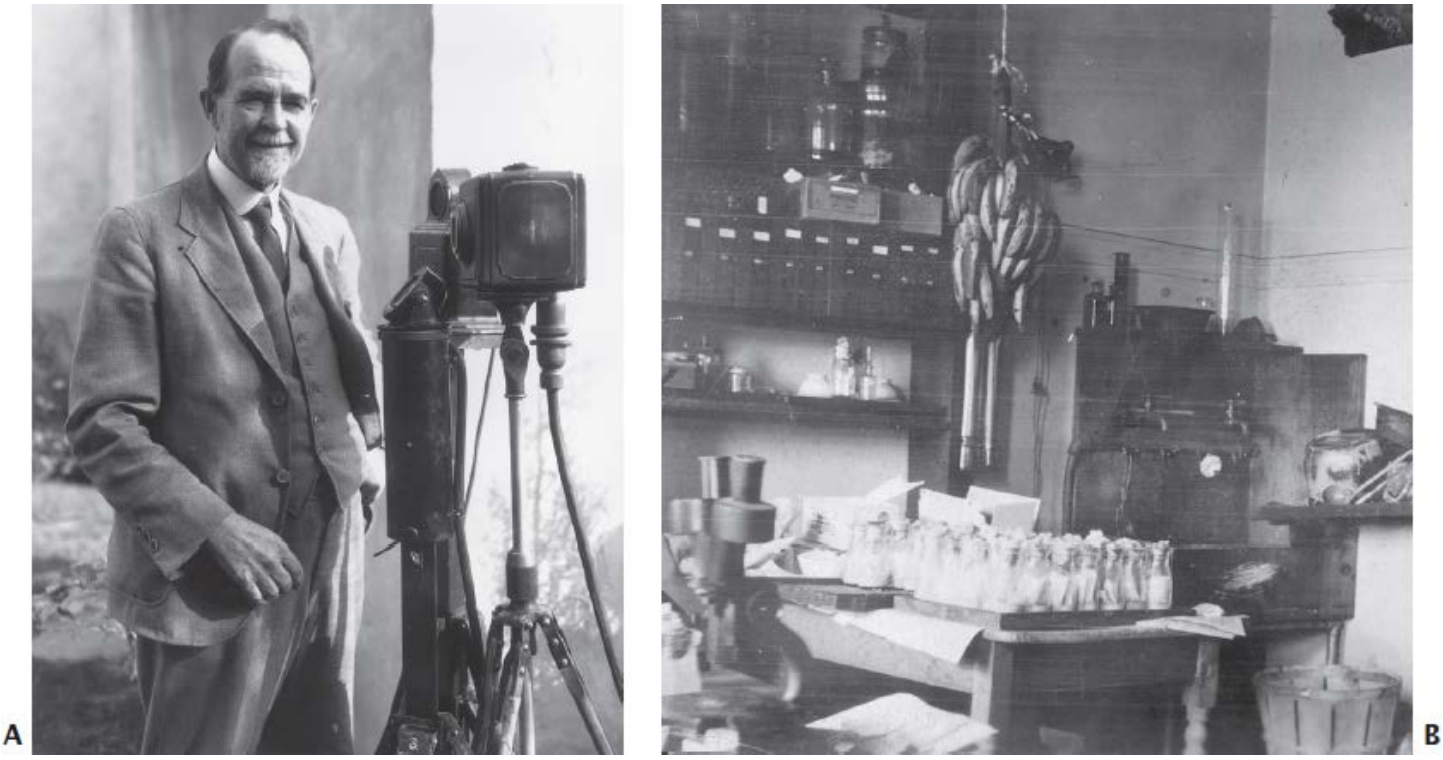


Figura 4.9 O trabalho de Thomas Hunt Morgan com a *Drosophila* ajudou a revelar muitos princípios básicos da genética, incluindo a herança ligada ao X. A. Morgan. **B.** O Quarto da Mosca, onde Morgan e seus estudantes fizeram a pesquisa genética. (*Parte A*: AP/Wide World Photos. *Parte B*: American Philosophical Society.)

Experimento

Pergunta: Os olhos brancos nas moscas-da-fruta são herdados como um traço recessivo autossômico?

Métodos Fazer cruzamentos recíprocos.

A Fêmea de olhos vermelhos cruzou com macho de olhos brancos

Geração P

Fêmea de olhos vermelhos × Macho de olhos brancos

X^+X^+ × X^wY

Meiose

Gametas: X^+ ; X^w , Y

Fertilização

Geração F_1

Fêmea de olhos vermelhos × Macho de olhos vermelhos

B Fêmea de olhos brancos cruzou com macho de olhos vermelhos

Geração P

Fêmea de olhos brancos × Macho de olhos vermelhos

X^wX^w × X^+Y

Meiose

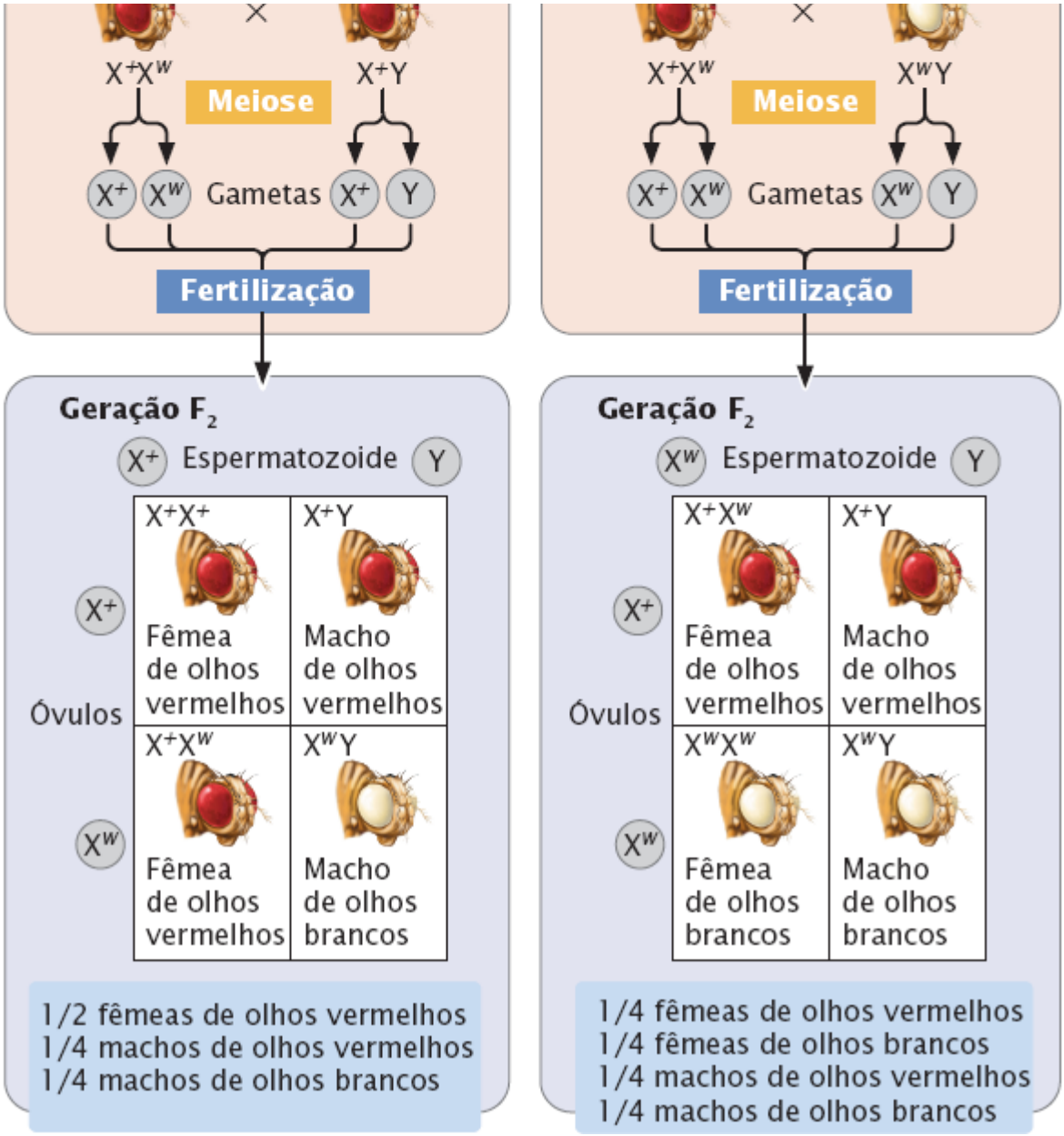
Gametas: X^w ; X^+ , Y

Fertilização

Geração F_1

Fêmea de olhos vermelhos × Macho de olhos brancos

Resultados



Conclusão: Não. Os resultados dos cruzamentos recíprocos são consistentes com herança ligada ao X.

Figura 4.10 Os cruzamentos ligados ao X realizados por Morgan para os olhos brancos nas moscas-da-fruta. A. Cruzamentos original e F_1. **B.** Cruzamentos recíprocos.

Para explicar esse resultado inesperado, Morgan propôs que o *locus* que afeta a cor do olho está no cromossomo X, (ou seja, a cor do olho é ligada ao X). Ele identificou que os alelos da cor dos olhos existem apenas no cromossomo X; não existe alelo homólogo no cromossomo Y. Como as fêmeas têm dois cromossomos X, elas podem ser homozigotas ou heterozigotas para os alelos da cor dos olhos. As células dos machos, por outro lado, têm apenas um cromossomo X e carreiam apenas um único alelo para esta característica. Portanto, os machos não podem ser homozigotos nem heterozigotos, e sim hemizigotos para os *loci* ligados ao X.

Para verificar sua hipótese de que o traço dos olhos brancos está ligado ao X, Morgan realizou cruzamentos adicionais. Ele previu que um cruzamento entre uma fêmea de olhos brancos e um macho de olhos vermelhos produziria todas as fêmeas de olhos vermelhos e todos os machos de olhos brancos (**Figura 4.10 B**). Quando Morgan fez este cruzamento, os resultados obtidos foram exatamente os previstos. Observe que este cruzamento é recíproco do cruzamento original e que os dois cruzamentos recíprocos produziram diferentes resultados nas gerações F_1 e F_2. Morgan também cruzou as fêmeas heterozigotas de F_1 com seu genitor de olhos brancos, as fêmeas de olhos vermelhos de F_2 com os machos de olhos brancos e as fêmeas de olhos brancos com os machos de olhos brancos. Em todos os cruzamentos, os resultados foram consistentes com a conclusão de Morgan de que o traço de olhos brancos é uma característica ligada ao X.

Não disjunção e teoria cromossômica da hereditariedade

Quando Morgan cruzou o macho original de olhos brancos com as fêmeas homozigotas de olhos vermelhos, todos os 1.237 descendentes tinham olhos vermelhos, exceto três machos. Morgan atribuiu estes machos de olhos brancos de F_1 à ocorrência de mutações aleatórias. Entretanto, as moscas com estes fenótipos inesperados continuavam a surgir nos seus cruzamentos. Embora incomum, elas apareciam com frequência excessiva para serem uma mutação espontânea. Calvin Bridges, um dos estudantes de Morgan, começou a investigar a base genética destas exceções.

Bridges descobriu que as exceções surgiam apenas em algumas cepas de moscas de olhos brancos. Quando ele cruzou uma destas fêmeas incomuns de olhos brancos com um macho de olhos vermelhos, cerca de 5% dos machos da geração tinham olhos vermelhos e cerca de 5% das fêmeas tinham olhos brancos. Neste cruzamento, o resultado esperado era de que todos os machos deviam herdar o cromossomo X da sua genitora e ter o genótipo X^w Y e olhos brancos (ver F_1 na Figura 4.10 B). Todas as moscas-fêmeas deveriam herdar um alelo dominante dos olhos vermelhos no cromossomo X do seu genitor, junto com um alelo para olhos brancos no cromossomo X de sua genitora; assim, todas as fêmeas descendentes devem ser X^+ X^w e ter olhos vermelhos (ver os descendentes de F_1 na Figura 4.10 B). O contínuo surgimento de machos de olhos vermelhos e fêmeas de olhos brancos neste cruzamento foi inesperado.

A explicação de Bridges. Para explicar o surgimento de machos de olhos vermelhos e fêmeas de olhos brancos neste cruzamento, Bridges criou a hipótese de que as raras fêmeas de olhos brancos desta cepa de fato tinham dois cromossomos X e um cromossomo Y (X^w X^w Y). Na *Drosophila*, as moscas com os cromossomos sexuais XXY normalmente se desenvolvem como fêmeas, apesar de terem um cromossomo Y (ver o Quadro 4.2). Em torno de 90% das vezes, os dois cromossomos X das fêmeas X^w X^w Y se separam na anáfase I da meiose, com um cromossomo X e um cromossomo Y entrando em um gameta e um único X entrando no outro gameta (**Figura 4.11**). Quando estes gametas são fertilizados pelos espermatozoides de um macho normal de olhos vermelhos, são produzidos machos de olhos brancos e fêmeas de olhos vermelhos. Em torno de 10% das vezes, os dois cromossomos X na fêmea não conseguem se separar na anáfase I da meiose, um fenômeno conhecido como **não disjunção**. Quando ocorre a não disjunção de X, metade dos óvulos recebe duas cópias do cromossomo X e a outra metade recebe apenas um cromossomo Y (ver a Figura 4.11). Quando estes óvulos são fertilizados pelos espermatozoides de um macho normal de olhos vermelhos, são produzidas quatro combinações de cromossomos sexuais. Um óvulo com dois cromossomos X que é fertilizado por um espermatozoide com X produz um zigoto X^+ X^w X^w, que geralmente morre. Quando um óvulo com dois cromossomos X é fertilizado por um esperma com cromossomo Y, o zigoto resultante é X^w X^w Y, que gera uma fêmea de olhos brancos. Um óvulo com apenas um cromossomo Y que é fertilizado por um espermatozoide com X produz um zigoto X^+ Y, que gera um macho normal com olhos vermelhos. Se o óvulo com apenas um cromossomo Y é fertilizado por um espermatozoide Y, o zigoto resultante tem dois cromossomos Y e nenhum cromossomo X e morre. A não disjunção dos cromossomos entre as fêmeas X^w X^w Y de olhos brancos, portanto, produz algumas fêmeas de olhos brancos e machos de olhos vermelhos, exatamente o que Bridges encontrou nos seus cruzamentos.

Confirmação da hipótese de Bridges. A hipótese de Bridges previa que as fêmeas de olhos brancos a partir desses cruzamentos teriam dois cromossomos X e Y e que alguns machos de olhos vermelhos teriam apenas um único cromossomo X. Para verificar sua hipótese, Bridges examinou os cromossomos de suas moscas e descobriu exatamente o que ele tinha previsto. A importância do estudo de Bridges não é de que tenha explicado o surgimento de uma mosca estranha na sua cultura, mas que ele foi capaz de relacionar a herança de um gene específico (*w*) com um cromossomo específico (X). Esta associação entre o genótipo e os cromossomos forneceu uma evidência clara de que os genes ligados ao sexo estão localizados no cromossomo X e confirmou a teoria cromossômica da herança. Resolva o Problema 22

Conceitos

Ao mostrar que o surgimento de fenótipos raros está associado à herança de cromossomos específicos, Bridges provou que os genes ligados ao sexo estão localizados no cromossomo X e que a teoria cromossômica da herança está correta.

Checagem dos conceitos 5

Qual era o genótipo dos machos de olhos vermelhos de F_1 obtidos por Bridges quando ele cruzou uma fêmea de olhos brancos com um macho de olhos vermelhos?

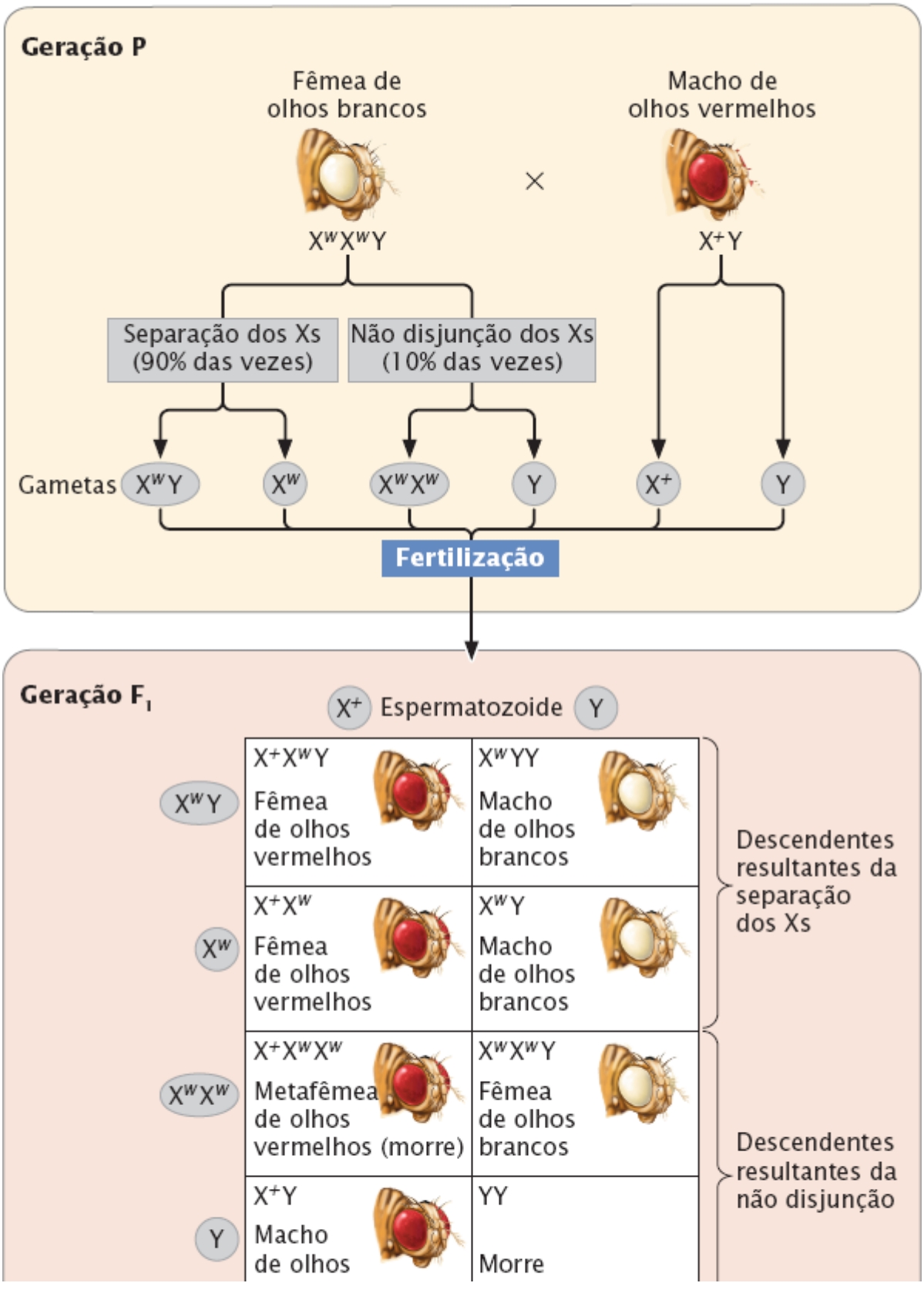
a. X^+.
b. X^wX^+Y.
c. X^+Y.
d. X^+X^+Y.

Experimento

Pergunta: Em um cruzamento entre uma fêmea de olhos brancos e um macho de olhos vermelhos, por que são produzidas fêmeas de olhos brancos e machos de olhos vermelhos?

Métodos

Hipótese: As fêmeas de olhos brancos e os machos de olhos vermelhos em F_1 resultam de não disjunção em uma fêmea XXY.



vermelhos

Resultados

Moscas de F_1	Cromossomos observados
Fêmeas de olhos vermelhos	1/2 XX, 1/2 XXY
Machos de olhos brancos	1/2 XY, 1/2 XYY
Fêmeas de olhos brancos	XXY
Machos de olhos vermelhos	XY

Conclusão: As fêmeas de olhos brancos e os machos de olhos vermelhos em F_1 resultam da não disjunção dos cromossomos X em uma fêmea XXY.

Figura 4.11 Bridges fez experimentos que provaram que o gene para os olhos brancos está localizado no cromossomo X.

Daltonismo ligado ao X nos seres humanos

Para examinar a herança ligada ao X, vamos considerar outra característica ligada a esse cromossomo: o daltonismo das cores vermelho-verde nos seres humanos. O olho humano percebe a cor por meio de cones sensíveis à luz que revestem a retina. Cada célula desse tipo tem um dos três pigmentos capazes de absorver a luz de um comprimento de onda específico: um absorve a luz azul, um segundo absorve a vermelha e o terceiro, a verde. Na prática, o olho humano detecta apenas três cores – vermelho, verde e azul – mas o cérebro mistura os sinais oriundos de diferentes cones para criar o amplo espectro de cores que percebemos. Cada um dos três pigmentos é codificado por um *locus* separado; o *locus* para o pigmento azul é encontrado no cromossomo 7, e os que codificam os pigmentos verde e vermelho estão juntos no cromossomo X.

Os tipos mais comuns de discromatopsia por cor nos seres humanos são causados por defeitos nos pigmentos vermelho e verde; vamos chamar essas condições de daltonismo das cores vermelho-verde. As mutações que comprometem a visão de cores são, em geral, recessivas, porque os genes que codificam os pigmentos vermelho e verde estão localizados no cromossomo X; o daltonismo das cores vermelho-verde é herdado como um traço recessivo ligado ao X.

Usaremos o símbolo X^c para representar um alelo para o daltonismo das cores vermelho-verde e o símbolo X^+ para representar um alelo para a visão de cor normal. As mulheres têm dois cromossomos X; então existem três possíveis genótipos entre elas: X^+X^+ e X^+X^c, que produzem a visão normal, e X^cX^c, que produz o daltonismo. Os homens têm apenas um único cromossomo X e dois possíveis genótipos: X^+ Y que produz a visão normal e X^cY que produz o daltonismo.

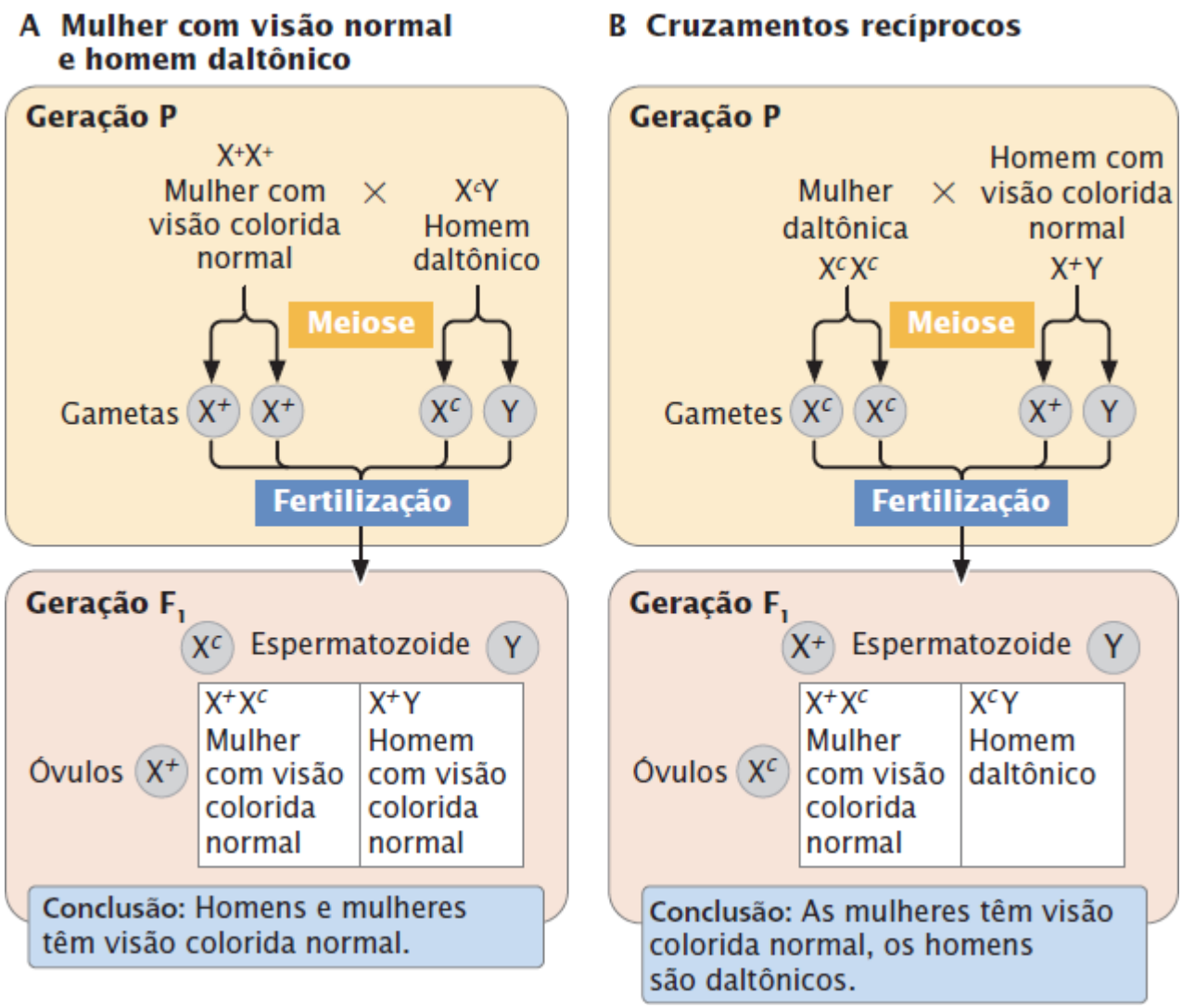


Figura 4.12 O daltonismo das cores vermelho-verde é herdado como um traço recessivo ligado ao X nos seres humanos.

Se uma mulher homozigota para visão normal tem relações sexuais com um homem daltônico (**Figura 4.12 A**), todos os gametas produzidos pela mulher terão um alelo para a visão colorida normal. Metade dos gametas do homem receberá o cromossomo X com o alelo para daltonismo e a outra metade receberá o cromossomo Y, que não tem alelos que afetem a visão colorida. Quando um espermatozoide com X^c se une a um óvulo com X^+, é produzida uma mulher heterozigota com visão normal ($X^+ X^c$). Quando um espermatozoide com Y se une a um óvulo com X, é produzido um homem heterozigoto com visão normal ($X^+ Y$).

No cruzamento recíproco entre uma mulher daltônica e um homem com visão colorida normal (**Figura 4.12 B**), a mulher produz apenas gametas com X^c. O homem produz alguns gametas que têm o cromossomo X e outros com cromossomo Y. Os homens herdam o cromossomo X das suas mães e, como ambos os cromossomos X da mãe têm o alelo X^c, todos os filhos serão daltônicos. Ao contrário, as mulheres herdam um cromossomo X dos pais; assim todas as filhas deste cruzamento recíproco serão heterozigotas com visão normal. As mulheres são daltônicas apenas quando os alelos para daltonismo são herdados de ambos os pais, enquanto um homem daltônico precisa herdar apenas um alelo do daltonismo da sua mãe; por isso, o daltonismo e outras características recessivas mais raras ligadas ao X são mais comuns nos homens.

Nesses cruzamentos para daltonismo, observe que uma mulher daltônica passa o traço recessivo ligado ao X para seus filhos, mas não para suas filhas, enquanto um homem daltônico passa o traço para seus netos por intermédio de suas filhas, através de suas filhas, mas nunca para seus filhos. As características recessivas ligadas ao X podem alternar-se entre os sexos, aparecendo nas mulheres em uma geração e nos homens na próxima geração.

Lembre-se de que os cromossomos X e Y pareiam na meiose porque eles são homólogos nas pequenas regiões pseudoautossômicas. Os genes nessas regiões do cromossomo X e Y são homólogos, como nos autossomos, e exibem padrões autossômicos de herança em vez da herança ligada ao sexo observada na maior parte dos genes nos cromossomos X e Y.

Problema desenvolvido

Agora que compreendemos o padrão da herança ligada ao X, vamos aplicar nosso conhecimento para responder uma

pergunta específica.

Betty tem a visão normal, mas sua mãe é daltônica. Bill é daltônico. Se Bill e Betty se casarem e tiverem um filho, qual é a probabilidade de esta criança ser daltônica?

Solução

De quais informações você precisa para solucionar o problema?

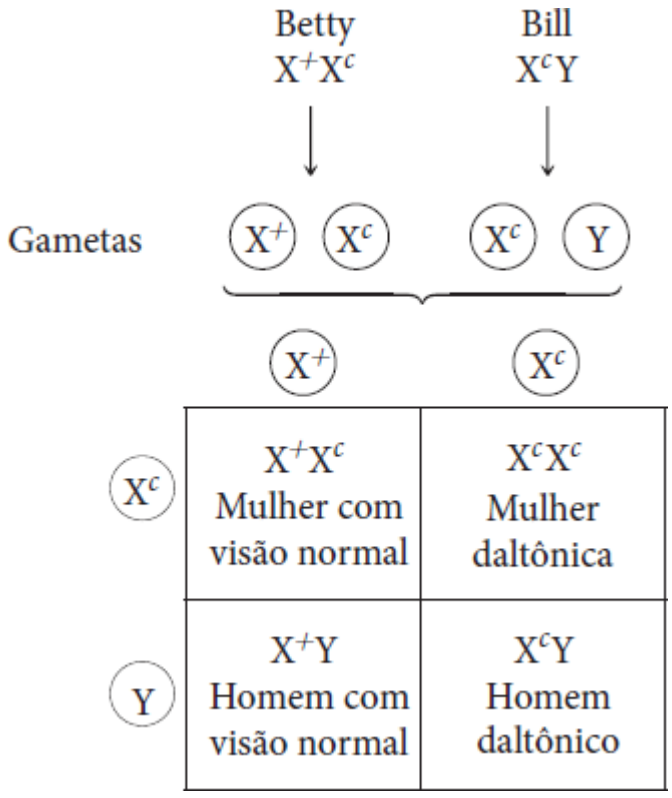
A probabilidade de que o filho de Bill e Betty seja daltônico.

Quais informações são fornecidas para solucionar o problema?

Os fenótipos de Betty, da mãe de Betty e de Bill.

Etapas para a solução

Como o daltonismo é um traço recessivo ligado ao X, a mãe de Betty, que é daltônica, tem de ser homozigota para o alelo do daltonismo (X^c X^c). As mulheres herdam um cromossomo X do pai e da mãe, então Betty tem de ter herdado um alelo do daltonismo de sua mãe. Como Betty tem a visão colorida normal, ela necessariamente herdou um alelo para visão normal (X^+) de seu pai, então Betty é heterozigota (X^+ X^c). Bill é daltônico. Como os homens são hemizigotos para os alelos ligados ao X, ele tem de ser (X^cY). Um cruzamento entre Betty e Bill é representado como:



Assim, a probabilidade de o filho seja daltônico é de 1/2.

Resolva o **Problema 24** no final deste capítulo para adquirir mais prática sobre a herança ligada ao X.

Conceitos

As características determinadas pelos genes nos cromossomos sexuais são chamadas de características ligadas ao sexo. As fêmeas diploides têm dois alelos em cada *locus* ligado ao X, enquanto os machos diploides têm um único alelo em cada *locus* ligado ao X. As fêmeas herdam os alelos ligados ao X dos genitores, mas os machos herdam um único alelo ligado ao X de suas genitoras.

Checagem dos conceitos 6

A hemofilia é uma doença recessiva ligada ao X nos seres humanos. Uma mulher com hemofilia tem relações sexuais com um homem com coagulação sanguínea normal. Qual a probabilidade de seu filho ter hemofilia?

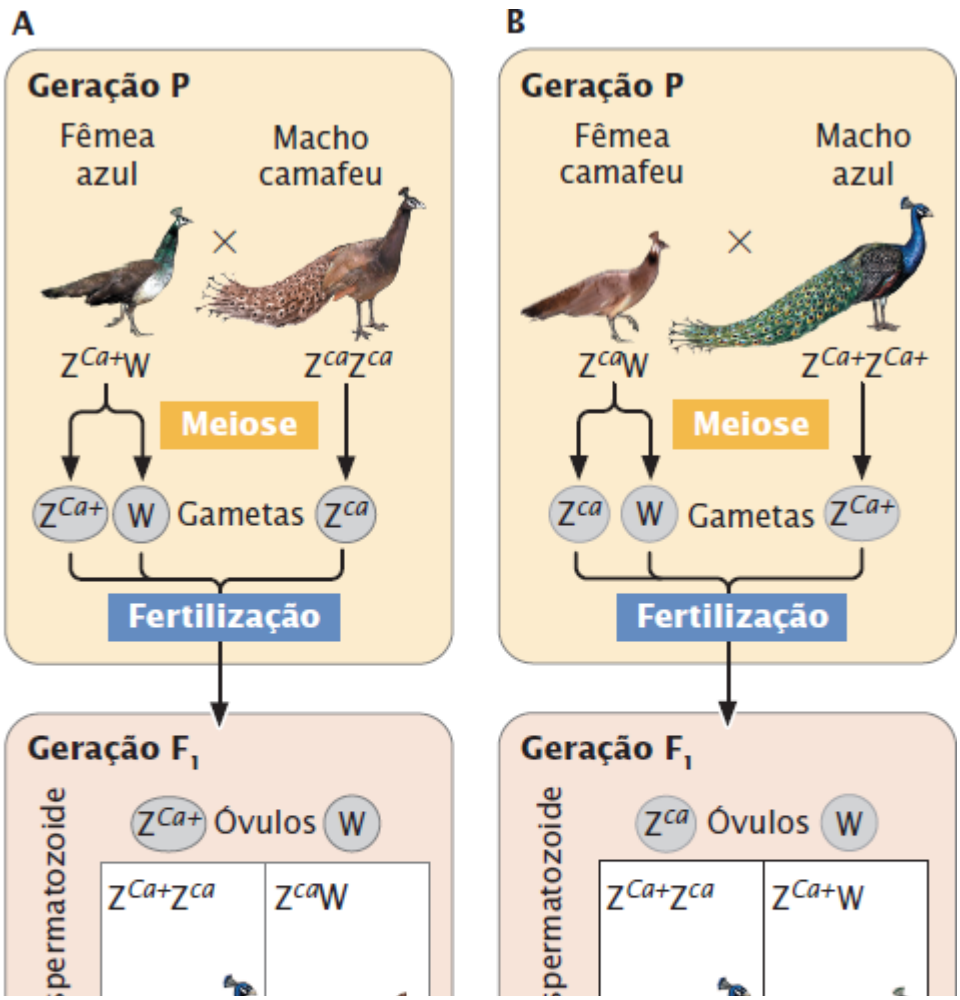
Símbolos para os genes ligados ao X

Existem várias formas diferentes para registrar os genótipos para os traços ligados ao X. Em alguns casos, eles são registrados da mesma forma que as características autossômicas. Assim, os machos hemizigotos recebem apenas um único alelo: por exemplo, o genótipo da *Drosophila* fêmea com olhos brancos é *ww* e o genótipo de um macho hemizigoto com olhos brancos é *w*. Outro método é incluir o cromossomo Y, indicando-o com uma barra diagonal (*/*). Com este método, o genótipo da fêmea de olhos brancos ainda é *ww* e o genótipo do macho de olhos brancos é *w/*. Talvez o método mais útil seja escrever os cromossomos X e Y no genótipo, indicando os alelos ligados ao X como sobrescritos, como é feito neste capítulo. Com este método, a fêmea de olhos brancos é X^wX^w e o macho de olhos brancos é X^wY. O uso de Xs e Ys no genótipo tem a vantagem de nos lembrar de que os genes são ligados ao X e de que os machos devem sempre ter um único alelo do tipo, herdado de sua genitora.

Características ligadas ao Z

Nos organismos com determinação do sexo pelo sistema ZZ-ZW, os machos são o sexo homogamético (ZZ) e carreiam dois alelos ligados ao sexo (em geral chamados de ligados ao Z), então eles podem ser homozigotos ou heterozigotos. As fêmeas são o sexo heterogamético (ZW) e têm apenas um único alelo ligado ao Z. A herança das características ligadas ao Z é a mesma das ligadas ao X, exceto que o padrão de herança em machos e fêmeas é invertido.

Exemplo de uma característica ligada ao Z é o fenótipo do camafeu no pavão-indiano azul (*Pavo cristatus*). Nessas aves, a plumagem selvagem é um azul metálico brilhoso. O pavão-fêmea é ZW e o macho é ZZ. A plumagem do camafeu, que produz as penas marrons, é resultado de um alelo ligado ao Z (Z^{ca}) que é recessivo para o alelo azul selvagem (Z^{Ca+}). Se uma fêmea de cor azul ($Z^{Ca+}W$) cruza com um macho camafeu ($Z^{ca}\ Z^{ca}$), todas as fêmeas de F_1 são camafeus ($Z^{ca}\ W$) e todos os machos de F_1 são azuis ($Z^{Ca+}Z^{ca}$), como demonstrado na **Figura 4.13**. Quando os descendentes de F_1 são intercruzados, 1/4 de F_2 são machos azuis ($Z^{Ca+}Z^{ca}$), 1/4 são fêmeas azuis ($Z^{Ca+}W$), 1/4 são machos camafeus ($Z^{ca}\ Z^{ca}$) e 1/4 são fêmeas camafeus ($Z^{ca}\ W$). O cruzamento recíproco de uma fêmea camafeu com um macho azul homozigoto produz uma geração F_1 na qual todos os descendentes são azuis e uma geração F_2 com 1/2 de machos azuis ($Z^{Ca+}Z^{Ca+}$ e $Z^{Ca+}Z^{ca}$), 1/4 de fêmeas azuis ($Z^{Ca+}W$) e 1/4 de fêmeas camafeus ($Z^{ca}\ W$).



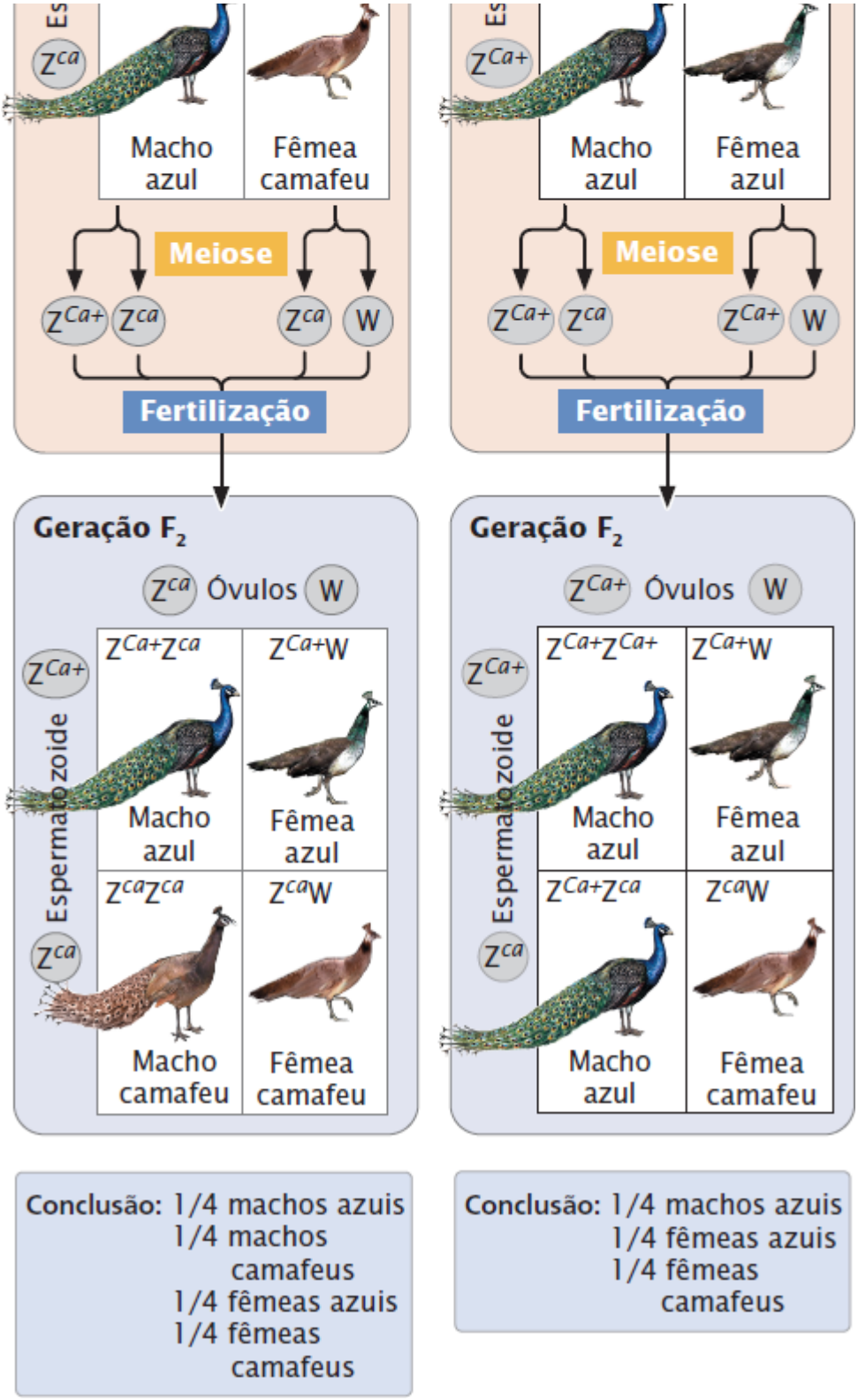


Figura 4.13 O fenótipo camafeu no pavão-indiano azul é herdado como um traço recessivo ligado ao Z. A. A fêmea azul cruzou com o macho camafeu. **B.** O cruzamento recíproco da fêmea camafeu que cruzou com o macho azul homozigoto.

Nos organismos com determinação do sexo pelo sistema ZZ-ZW, a fêmea sempre herda seu cromossomo W de sua genitora e seu cromossomo Z, junto com qualquer alelo ligado ao Z, de seu genitor. Neste sistema, o macho herda os cromossomos Z, junto com qualquer alelo ligado ao Z, de seus genitores. Este padrão de herança é o oposto do padrão dos alelos ligados ao X nos organismos com determinação do sexo pelo sistema XX-XY. Resolva o Problema 33

Características ligadas ao Y

Os traços ligados ao Y – também chamados de traços holândricos – exibem um padrão diferente de herança. Eles ocorrem apenas nos machos, porque apenas os machos têm um cromossomo Y, que é sempre herdado do pai. Além disso, todos os descendentes do sexo masculino de um macho com um traço ligado ao Y apresentarão este traço, porque todos os machos herdam o cromossomo Y do seu genitor. Resolva o Problema 45

Evolução do cromossomo Y. A pesquisa sobre os cromossomos sexuais levou à conclusão de que os cromossomos X e Y, em muitos organismos, evoluíram a partir de um par de autossomos. A primeira etapa neste processo evolutivo ocorreu quando um membro de um par de autossomos adquiriu um gene que determinava a masculinidade, como o gene *SRY* encontrado nos seres humanos atualmente (**Figura 4.14**). Esta etapa ocorreu nos mamíferos há 250 milhões de anos. Um organismo com uma cópia do cromossomo contendo esse gene então se tornou um macho. Ocorreram mutações adicionais no cromossomo proto-Y que afetaram os traços que eram benéficos apenas para os machos, como a coloração brilhante que as aves machos usam para atrair as fêmeas e os cornos que os alces usam na competição com outros machos. Os genes que codificam esses tipos de traços são vantajosos apenas se estiverem presentes nos machos. Para evitar que os genes que codificam os traços masculinos apareçam nas fêmeas, o *crossing over* foi suprimido para a maior parte da extensão dos cromossomos X e Y na meiose. Ele ainda pode ocorrer entre dois cromossomos X nas fêmeas, com pouca frequência entre os cromossomos X e Y, exceto para pequenas regiões pseudoautossômicas, nas quais os cromossomos X e Y continuam a parear e se separar na meiose, como discutido anteriormente.

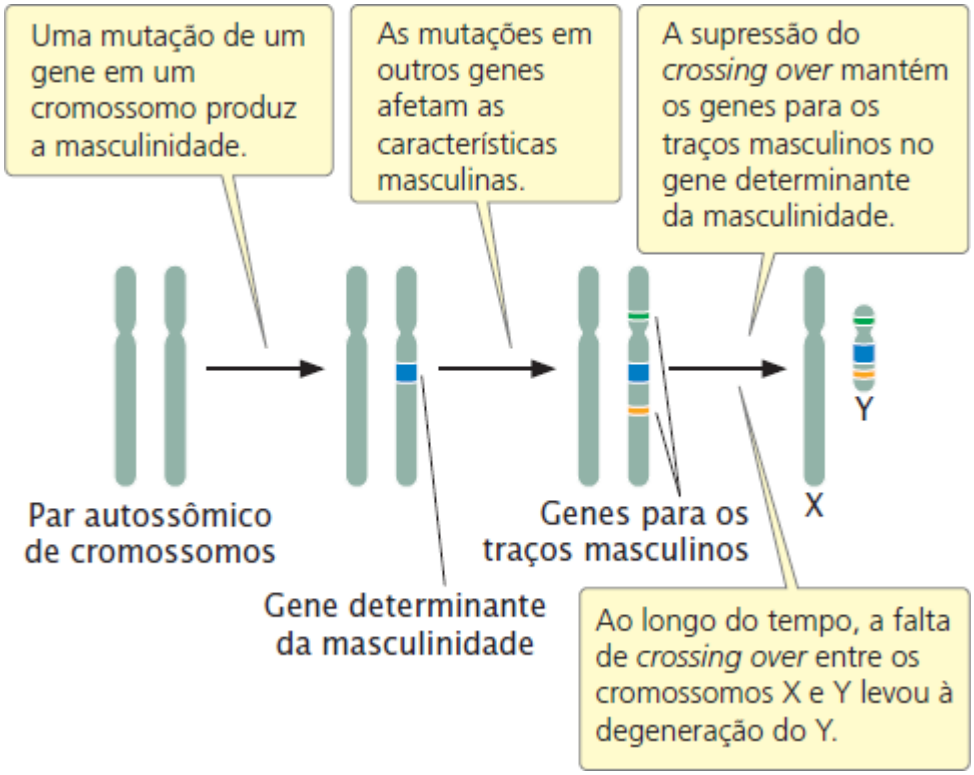


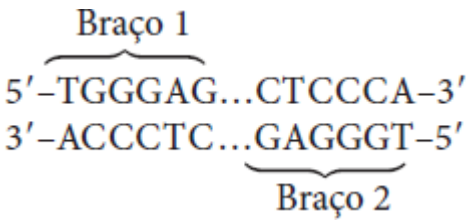
Figura 4.14 **Evolução do cromossomo Y.**

Por motivos que estão além do escopo da discussão aqui, a falta de *crossing over* leva (e continua a levar) a acúmulo de mutações e perda de material genético do cromossomo Y (ver Figura 4.14). Durante milhares de anos, o cromossomo Y lentamente se degenerou, perdendo DNA e genes até se tornar muito reduzido e reter pouca informação genética, produzindo o cromossomo Y encontrado nos machos atualmente. Na verdade, os cromossomos Y nos seres humanos e em muitos outros organismos são pequenos e têm pouca informação genética; portanto, poucas características exibem a herança ligada ao Y. Alguns pesquisadores previram que o cromossomo Y humano continuará a perder informação genética no futuro e vai desaparecer por completo da espécie em cerca de 10 milhões de anos, um panorama desanimador para os que têm um cromossomo Y (e talvez algumas com dois Xs). Entretanto, uma pesquisa publicada em 2012 sugere que a decomposição do cromossomo Y humano parou e não foram perdidos genes nos últimos 6 milhões de anos. A recombinação interna dentro do cromossomo Y (próxima Seção) pode ter ajudado a reduzir ou evitar a decomposição do cromossomo Y humano.

Características do cromossomo Y humano. A sequência genética da maior parte do cromossomo Y humano foi determinada recentemente como parte do Projeto do Genoma Humano (ver *Capítulo 19*). Esse trabalho revela que cerca de dois terços do cromossomo Y consiste em pequenas sequências de DNA que se repetem várias vezes (ver *Capítulo 9*) e não têm genes ativos. O outro terço tem apenas alguns genes. Apenas cerca de 350 genes foram identificados no cromossomo Y humano, comparados com os milhares na maior parte dos cromossomos, e apenas em torno de metade deles codificam proteínas. A função da maior parte dos genes ligados ao Y é pouco compreendida, parecendo que influenciam o desenvolvimento sexual e a fertilidade masculina. Alguns são expressos por todo o corpo, mas muitos são expressos

predominante ou exclusivamente nos testículos. Embora o cromossomo Y tenha relativamente poucos genes, a pesquisa recente na *Drosophila* sugere que ele carrega elementos genéticos que afetam a expressão de vários genes nos cromossomos autossômicos e no X.

Uma característica surpreendente revelada pelo sequenciamento é a existência de oito grandes sequências palindrômicas no cromossomo Y. Um palíndromo é definido como uma palavra ou frase em que se lê o mesmo da direita para a esquerda ou vice-versa. Uma sequência palindrômica no DNA lê a mesma informação em ambas as fitas da dupla-hélice, criando duas cópias quase idênticas que se alongam a partir de um ponto central, como:



Desse modo, uma sequência palindrômica de DNA aparece duas vezes, muito parecida com as duas cópias de uma sequência de DNA encontradas em dois cromossomos homólogos. Na verdade, a recombinação ocorre entre as duas sequências palindrômicas no cromossomo Y. Como já foi citado, os cromossomos X e Y não são homólogos em quase todas as suas sequências e a maior parte do cromossomo Y não sofre *crossing over* com o cromossomo X. Essa falta de recombinação intercromossômica leva a um acúmulo de mutações deletérias no cromossomo Y e à perda de material genético. As evidências sugerem que os dois braços do cromossomo Y se recombinam um com outro, o que, em parte, compensa a ausência de recombinação entre os cromossomos X e Y. Essa recombinação interna ajuda a manter algumas sequências e funções dos genes no cromossomo Y e evita sua total degeneração.

Embora os palíndromos ofereçam oportunidades para recombinação, o que ajuda a evitar a degradação do cromossomo Y dentro da escala evolutiva, elas têm efeitos prejudiciais. Pesquisa recente revelou que a recombinação entre os palíndromos pode levar a rearranjos do cromossomo Y que podem causar anomalias no desenvolvimento sexual. Em alguns casos, a recombinação entre os palíndromos leva à deleção do gene *SRY*, produzindo uma fêmea XY. Em outros casos, a recombinação elimina outros genes do cromossomo Y que participam na produção do espermatozoide. Em alguns casos, a recombinação produz um cromossomo Y com dois centrômeros, que se rompem à medida que os eles são puxados nas direções opostas na mitose. Os cromossomos Y rompidos podem ser perdidos na mitose, resultando em células XO e na síndrome de Turner.

Conceitos

Os traços ligados ao Y exibem um padrão diferente de herança: existem apenas nos machos e todos os descendentes do sexo masculino de um macho com um traço ligado ao Y herdam este traço. As sequências palindrômicas dentro do cromossomo Y podem sofrer recombinação interna, mas tal recombinação pode levar a anomalias cromossômicas.

Checagem dos conceitos 7

Qual característica rara do cromossomo Y permite recombinação entre os genes encontrados neste cromossomo?

Uso de marcadores genéticos ligados ao Y. As sequências de DNA no cromossomo Y sofrem mutação com o passar do tempo e variam entre os indivíduos do sexo masculino. Essas mutações criam variações na sequência de DNA – chamados de marcadores genéticos – que, como os traços ligados ao Y, são transmitidos do pai para o filho e podem ser usados para estudar a ascendência masculina. Embora os próprios marcadores não codifiquem os traços físicos, eles podem ser detectados com o uso de métodos moleculares. Boa parte do cromossomo Y não tem função, então as mutações se acumulam com facilidade. Muitas dessas mutações são únicas, surgem apenas uma vez e são transmitidas através das gerações. Os machos individuais com o mesmo conjunto de mutações estão relacionados uns com os outros, e a distribuição desses marcadores genéticos nos cromossomos Y atualmente fornece pistas sobre as relações genéticas das pessoas.

Os marcadores ligados ao Y têm sido usados para estudar os descendentes de Thomas Jefferson, principal autor da Declaração de Independência e terceiro presidente dos EUA. Em 1802, um inimigo político acusou Thomas Jefferson de ser o pai de um filho de sua escrava Sally Hemings, mas a evidência era circunstancial. Sally, que trabalhou para a família de Thomas e o acompanhou em uma viagem a Paris, teve cinco filhos. Jefferson foi acusado de ser o pai da primeira criança, mas rumores sobre a paternidade de outras crianças também circularam. Os antepassados dos filhos de Sally mantinham a teoria de que eram

descendentes de Thomas, mas alguns descendentes de Jefferson se recusavam a reconhecer sua reivindicação.

Para solucionar essa antiga controvérsia, os geneticistas examinaram marcadores dos cromossomos Y dos descendentes masculinos do primeiro filho de Hemings (Thomas Woodson), seu último filho (Eston Hemings) e um tio paterno de Thomas Jefferson com quem Jefferson tinha os cromossomos Y em comum – foram usados os descendentes do tio de Jefferson porque o próprio Jefferson não tinha descendentes masculinos. Os geneticistas determinaram que Jefferson tinha um conjunto raro e inconfundível de marcadores genéticos no seu cromossomo Y. Os mesmos marcadores também foram encontrados nos cromossomos Y dos descendentes do sexo masculino de Eston Hemings. A probabilidade de esta combinação ser aleatória é menor que 1%. Os marcadores não foram encontrados nos cromossomos Y dos descendentes de Thomas Woodson. Junto com os indícios históricos, estes marcadores combinados são mais sugestivos de que Jefferson era o pai de Eston Hemings, mas não de Thomas Woodson.

As sequências do cromossomo Y também foram extensivamente usadas para examinar padrões antigos de migração humana e as relações genéticas entre diferentes populações de humanos.

Conceitos conectantes

Como reconhecer a herança ligada ao sexo

Quais características devemos procurar para identificar um traço como ligado ao sexo? Uma concepção errada, mas comum, é de que qualquer característica genética na qual os fenótipos de homens e mulheres são diferentes tem de ser ligada ao sexo. Na verdade, a expressão de muitas características *autossômicas* é diferente entre machos e fêmeas. Os genes que codificam essas características são os mesmos em ambos os sexos, mas sua expressão é influenciada pelos hormônios sexuais, fazendo com que os mesmos genes produzam diferentes fenótipos nos homens e mulheres.

Outra ideia errada é a de que qualquer característica encontrada com maior frequência em um sexo está ligada ao sexo. Vários traços autossômicos são expressos mais comumente em um sexo do que em outro. Diz-se que esses traços são influenciados pelo sexo. Alguns traços autossômicos são expressos apenas em um sexo; eles são limitados ao sexo. Ambas as características influenciadas e limitadas ao sexo serão discutidas com mais detalhes no Capítulo 5.

Vários aspectos das características ligadas ao sexo facilitam sua identificação. Os traços ligados ao Y são encontrados apenas nos machos, mas isso não garante que eles sejam de fato ligados ao Y, porque algumas características autossômicas são expressas apenas nos machos. Um traço ligado ao Y é único, de modo que todos os descendentes do sexo masculino de um macho afetado expressem seu fenótipo e ele só possa ser herdado a partir do lado paterno da família. Assim, este traço pode ser herdado apenas do avô paterno, nunca do materno.

As características ligadas ao X também exibem um padrão diferente de herança. A ligação ao X é uma possível explicação quando os cruzamentos recíprocos têm resultados diferentes. Se uma característica está ligada ao X, um cruzamento entre um macho afetado e uma fêmea não afetada terá os mesmos resultados que um cruzamento entre uma fêmea afetada e um macho não afetado. Os cruzamentos recíprocos têm o mesmo resultado para quase todas as características autossômicas. No entanto, não devemos concluir que, quando os cruzamentos recíprocos têm resultados diferentes, a característica esteja ligada ao X. Outras formas de herança associadas ao sexo, discutidas no Capítulo 5, também produzem resultados diferentes nos cruzamentos recíprocos. A chave para identificar a herança ligada ao X é lembrar que um macho sempre herda seu cromossomo X da sua genitora, não do seu genitor. Assim, uma característica ligada ao X não é passada diretamente do pai para seu filho; se um macho herda um traço do pai e a mãe não é heterozigota, esse traço não pode estar ligado ao X.

4.3 A compensação da dosagem iguala a quantidade de proteína produzida por genes ligados ao X e genes autossômicos em alguns animais

Nas espécies com determinação do sexo pelo sistema XX-XY, as diferenças no número de cromossomos X presentes nos machos e nas fêmeas são uma questão especial no desenvolvimento. Nas fêmeas, existem duas cópias do cromossomo X e duas cópias de cada autossomo, então os genes nos cromossomos X e nos autossomos estão “equilibrados”. Nos machos, entretanto, existe apenas um cromossomo X, enquanto existem duas cópias de cada autossomo. Como a quantidade de proteína produzida

é, em muitos casos, dependente do número de cópias de genes que a codifica, é provável que nos machos exista menos proteína codificada pelos genes ligados ao X do que proteína codificada pelos genes autossômicos. Essa diferença pode ser prejudicial, porque a concentração de proteínas tem um papel crítico no desenvolvimento.

Alguns animais superaram esse problema ao desenvolver mecanismos para equilibrar a quantidade de proteína produzida por um único X e dois autossomos no sexo heterogamético, chamados de **compensação de dosagem**. Nas moscas-da-fruta, a compensação de dosagem é alcançada ao dobrar a atividade dos genes no cromossomo X dos machos, mas não nos das fêmeas. Nos mamíferos placentários, a expressão dos genes sensíveis à dose nos cromossomos X nos machos e fêmeas aumenta, acoplada com a inativação de um dos cromossomos X nas fêmeas, de modo que a expressão dos genes ligados ao X e autossômicos é equilibrada nos machos e nas fêmeas.

Por motivos que desconhecemos, a presença dos cromossomos sexuais nem sempre produz problemas na dose do gene e a compensação de dosagem dos genes ligados ao X não é universal. Vários animais não exibem mecanismos claros de compensação de dosagem, dentre ele borboletas e mariposas, aves e alguns peixes, e até o ornitorrinco. Como discutiremos na próxima Seção, até nos mamíferos placentários vários genes escapam da compensação de dosagem.

Hipótese de Lyon

Em 1949, Murray Barr observou corpúsculos condensados, escuros, nos núcleos das células de gatas (**Figura 4.15**); estas estruturas ficaram conhecidas como **corpúsculos de Barr**. Mary Lyon propôs, em 1961, que o corpúsculo de Barr era um cromossomo X inativo; sua hipótese (atualmente aceita para mamíferos placentários) ficou conhecida como a **hipótese de Lyon**. Ela sugeriu que, dentro de cada célula da fêmea, um dos dois cromossomos X se torna aleatoriamente inativo. Se uma célula tem mais de dois cromossomos X, todos exceto um estão inativos. O número de corpúsculos de Barr nas células humanas com diferentes complementos de cromossomos sexuais é apresentado no **Quadro 4.3**.

Como resultado da inativação do X, as fêmeas de mamíferos placentários são funcionalmente hemizigotas em nível celular para os genes ligados ao X. Nas fêmeas heterozigotas em um *locus* ligado ao X, aproximadamente 50% das células expressará um alelo e 50% expressará o outro alelo; assim, nas fêmeas heterozigotas, as proteínas codificadas por ambos os alelos são produzidas, embora não na mesma célula. Essa **hemizigosidade** funcional significa que as células nas fêmeas não são idênticas com relação à expressão dos genes no cromossomo X; as fêmeas são mosaicos para a expressão dos genes ligados ao X.

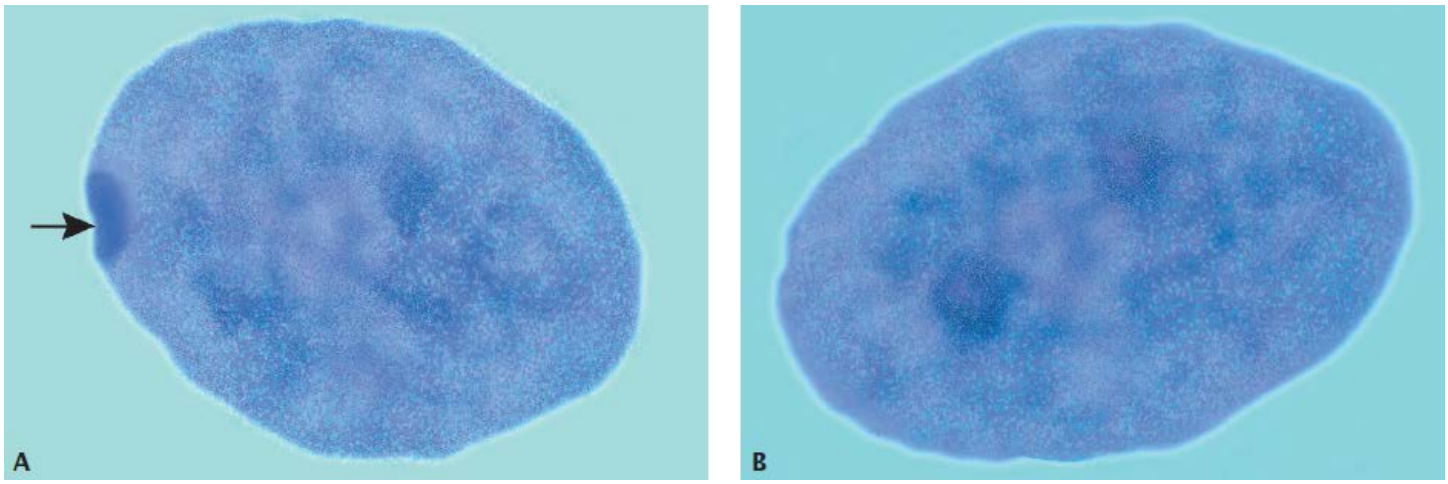


Figura 4.15 Um corpúsculo de Barr é um cromossomo X inativo. A. Célula de fêmea com um corpúsculo de Barr (*indicado pela seta*). **B.** Célula de macho sem corpúsculo de Barr. (*Chris Bjornberg/Photo Researchers.*)

Quadro 4.3 O número de corpúsculos de Barr existentes nas células humanas com diferentes complementos de cromossomos sexuais.

Cromossomos sexuais	Síndrome	Número de corpúsculos de Barr
XX	Nenhuma	1
XY	Nenhuma	0
XO	Turner	0
XXY	Klinefelter	1

XXYY	Klinefelter	1
XXXY	Klinefelter	2
XXXXY	Klinefelter	3
XXX	X Triplo	2
XXXX	Fêmea poli-X	3
XXXXX	Fêmea poli-X	4

A inativação aleatória de X ocorre no início do desenvolvimento – nos seres humanos, dentro das primeiras semanas de desenvolvimento. Após um cromossomo X ser inativado em uma célula, ele permanece assim em todas as células somáticas que descendem desta célula. Assim, as células vizinhas tendem a ter o mesmo cromossomo X inativo, produzindo um padrão irregular (mosaico) para a expressão de uma característica ligada ao X nas fêmeas heterozigotas.

Essa distribuição irregular pode ser observada nos gatos casco de tartaruga e cálicos (**Figura 4.16**). Embora muitos genes contribuam para a cor e o padrão do pelo nos gatos domésticos, um único *locus* ligado ao X determina a cor laranja. Existem dois possíveis alelos neste *locus*: X^+, que produz um pelo não laranja (em geral preto), e X^o, que produz o pelo laranja. Os machos são hemizigotos, então podem ser pretos (X^+ Y) ou laranja (X^o Y), mas não preto e laranja (podem surgir raros gatos casco de tartaruga machos a partir de dois cromossomos X, X^+ X^o Y). As fêmeas podem ser pretas (X^+ X^+), laranja (X^o X^o) ou casco de tartaruga (X^+ X^o). Cada segmento de pelo laranja é um clone de células derivadas de uma célula original na qual o alelo preto está inativo, e cada segmento de pelo preto é um clone de células derivadas de uma célula original na qual o alelo laranja está inativo.

Figura 4.16 A distribuição irregular de cor nos gatos casco de tartaruga resulta de inativação aleatória de um cromossomo X nas fêmeas. (Robert Adrian Hillman/Shutterstock.)

A hipótese de Lyon sugere que a existência de números variáveis de cromossomos X não deve afetar o fenótipo nos mamíferos, porque qualquer cromossomo X a mais deve ser inativo. Entretanto, as pessoas com a síndrome de Turner (XO) são diferentes das mulheres XX e as pessoas com síndrome de Klinefelter (XXY) são diferentes dos homens XY. Como estas condições surgem diante da compensação de dosagem?

É provável que esses distúrbios surjam porque alguns genes ligados ao X escapam da inativação. De fato, a natureza da inativação do X é mais complexa do que se imaginou originalmente. Estudos com genes individuais revelam que apenas cerca de 75% dos genes humanos ligados ao X ficam permanentemente inativos. Cerca de 15% escapam completamente da inativação do X, o que significa que eles produzem duas vezes mais proteínas nas mulheres que nos homens. Os 10% remanescentes estão inativos em algumas mulheres, mas não em todas. O motivo para esta variação entre as mulheres na ativação de alguns genes ligados ao X é desconhecido. Além disso, pesquisa recente indica que a inativação do X não equilibra realmente a dosagem de muitos genes ligados ao X e autossômicos nos seres humanos e nos camundongos. **Resolva o Problema 44**

Mecanismo de inativação aleatória do X

A inativação aleatória dos cromossomos X ocorre em duas etapas. Na primeira, a célula avalia, ou conta, de alguma forma quantos cromossomos X existem. Na segunda etapa, um cromossomo X é selecionado para se tornar o cromossomo ativo e todos os outros são desativados.

Embora ainda não sejam conhecidos muitos detalhes da inativação do cromossomo X, já foram identificados vários genes e sequências que participam do processo. O mais notável entre eles é um gene chamado de *Xist* (para transcrição específica de inativação do X, *X-inactivation-specific transcript*). Nos cromossomos X a serem inativados, o gene *Xist* está ativo, produzindo uma molécula de RNA com 17.000 nucleotídios que reveste o cromossomo X e inativa seus genes, provavelmente ao recrutar complexos de proteínas que alteram a estrutura da cromatina. No cromossomo X que se tornará ativo, outros genes reprimem a atividade do *Xist* de modo que o RNA nunca cubra o cromossomo e os genes permaneçam ativos.

Conceitos

Nos mamíferos placentários, um dos cromossomos X é inativados em cada célula; qual deles é inativado é aleatório e varia entre as células.

Checagem dos conceitos 8

Quantos corpúsculos de Barr um macho com cromossomos XXXYY terá em cada uma de suas células?

Resumo dos conceitos

- A reprodução sexuada é a produção de descendentes que são geneticamente diferentes de seus genitores. A maior parte dos organismos tem dois fenótipos sexuais: macho e fêmea. Os machos produzem gametas pequenos, as fêmeas produzem gametas grandes
- O mecanismo pelo qual o sexo é estabelecido é chamado de determinação do sexo. Ele pode ser determinado pelas diferenças em cromossomos específicos, pelos genótipos ou pelo meio ambiente
- Os cromossomos sexuais de machos e fêmeas têm aparência diferente e estão em diferente quantidade. O sexo homogamético produz gametas que são todos idênticos com relação aos cromossomos sexuais; o sexo heterogamético produz gametas com diferente composição dos cromossomos sexuais
- No sistema XX-XO de determinação do sexo, as fêmeas têm dois cromossomos X, enquanto os machos têm um único cromossomo X. No sistema XX-XY, as fêmeas têm dois cromossomos X, enquanto os machos têm um único cromossomo X e um cromossomo Y. No sistema ZZ-ZW, os machos têm dois cromossomos Z, enquanto as fêmeas têm um cromossomo Z e um W
- Alguns organismos têm a determinação do sexo pelo gene, na qual os genótipos em um ou mais *loci* determinam o sexo de um indivíduo. Mas ainda existem outros com a determinação do sexo pelo meio ambiente
- Na *Drosophila melanogaster*, o sexo é determinado pela razão X:A, mas é determinado principalmente pelos genes no cromossomo X
- Nos seres humanos, o sexo é determinado pela presença ou ausência do gene *SRY* localizado no cromossomo Y
- As características ligadas ao sexo são determinadas pelos genes presentes nos cromossomos sexuais: as características ligadas ao X são codificadas pelos genes no cromossomo X; as características ligadas ao Y são codificadas pelos genes no cromossomo Y
- Uma fêmea herda os alelos ligados ao X de ambos os genitores, um macho herda os alelos ligados ao X apenas de sua genitora
- Os cromossomos sexuais evoluíram a partir dos autossomos. O *crossing over* entre os cromossomos X e Y foi suprimido, mas as sequências palindrômicas no cromossomo Y possibilitam a recombinação interna nesse cromossomo. Esta recombinação interna às vezes leva a rearranjos no cromossomo que podem afetar de maneira adversa o desenvolvimento sexual
- As características ligadas ao Y são encontradas apenas nos machos e são passadas de pai para filho
- Nos mamíferos placentários, um dos dois cromossomos X nas fêmeas normalmente é inativado. Qual cromossomo X é inativado é aleatório e varia entre as células. Alguns genes ligados ao X escapam da inativação do X e outros podem ser inativos em algumas fêmeas, mas não em outras. A inativação do X é controlada pelo gene *Xist*.

Termos importantes

Autossomo
Características ligadas ao sexo

Características ligadas ao X
Características ligadas ao Y
Compensação de dosagem
Corpúsculo de Barr
Cromossomos sexuais
Determinação do sexo
Determinação do sexo pelo gene
Gene da região determinante do sexo no cromossomo Y (*SRY*)
Hemizigosidade
Hermafroditismo
Hermafroditismo sequencial
Hipótese de Lyon
Não disjunção
Organismo dioico
Organismo monoico
Regiões pseudoautossômicas
Sexo (p. 70)
Sexo heterogamético
Sexo homogamético
Síndrome de Klinefelter
Síndrome de Turner
Síndrome do X triplo

Respostas da Checagem dos conceitos

1. Meiose.
2. b.
3. Na determinação do sexo pelo cromossomo, machos e fêmeas têm cromossomos que são distinguíveis. Na determinação do sexo pelo gene, o sexo é determinado pelos genes, mas os cromossomos de machos e fêmeas são indistinguíveis. Na determinação do sexo pelo meio ambiente, o sexo é determinado em parte ou totalmente por fatores ambientais.
4. a.
5. c.
6. Todos os descendentes masculinos terão hemofilia e todos os femininos não a terão; então a probabilidade geral da hemofilia nos descendentes é de 1/2.
7. Oito grandes palíndromos permitem o *crossing over* no cromossomo Y.
8. Dois corpúsculos de Barr.

Problemas desenvolvidos

Problema 1

Uma mosca-da-fruta tem cromossomos sexuais XXXYY; todos os cromossomos autossômicos são normais. Qual é o fenótipo sexual desta mosca?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
O fenótipo sexual de uma mosca com cromossomos sexuais XXXYY.

Quais informações são fornecidas para solucionar o problema?

- A mosca tem cromossomos sexuais XXXYY.
- Todos os cromossomos autossômicos são normais.

Para a solução deste problema, revise:
Determinação do sexo na *Drosophila melanogaster*, na Seção 4.1.

Etapas para a solução

O sexo nas moscas-da-fruta é determinado pela razão X:A, a razão do número de cromossomos X pelo número de conjuntos haploides de cromossomos autossômicos. Uma razão X:A de 1,0 produz uma mosca-fêmea; uma razão X:A de 0,5 produz um macho. Se a razão X:A for maior que 1,0, a mosca é metafêmea, se for menor que 0,5, a mosca é metamacho, se a razão X:A estiver entre 1,0 e 0,5, a mosca é um intersexo.

Lembrete: A *Drosophila melanogaster* normalmente tem dois conjuntos de autossomos.

Essa mosca tem três cromossomos X e autossomos normais, então a proporção X:A neste caso é 3/2 ou 1,5. Então, esta mosca é uma metafêmea.

Problema 2

Na *Drosophila melanogaster*, as cerdas bifurcadas são causadas por um alelo (X^f) que é ligado ao X e recessivo para um alelo para cerdas normais (X^+). Os olhos marrons são produzidos por um alelo (*b*) autossômico e recessivo para um alelo para olhos vermelhos (b^+). Uma mosca-fêmea que é homozigota para cerdas normais e olhos vermelhos cruza com uma mosca-macho que tem cerdas bifurcadas e olhos marrons. Os descendentes F_1 são intercruzados para produzir F_2. Quais serão os fenótipos e as proporções das moscas de F_2 a partir deste cruzamento?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Os fenótipos e as proporções das moscas de F_2.

Quais informações são fornecidas para solucionar o problema?

- As cerdas bifurcadas são recessivas ligadas ao X
- Os olhos castanhos são autossômicos recessivos
- Os fenótipos dos genitores do cruzamento
- Os descendentes de F_1 são intercruzados para produzir F_2.

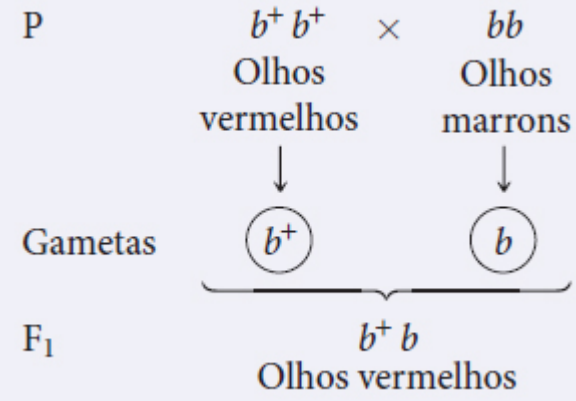
Para a solução deste problema, revise:
Daltonismo ligado ao X nos seres humanos, na Seção 4.2 e Seção 3.3, no Capítulo 3.

Etapas para a solução

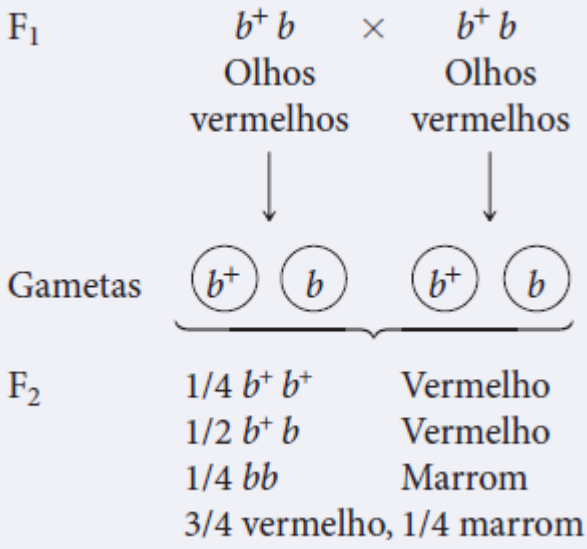
É melhor resolver este problema decompondo o cruzamento em dois, um para os genes ligados ao X que determinam o tipo de cerdas e um para os genes autossômicos que determinam a cor dos olhos.

Vamos começar com as características autossômicas. Uma mosca-fêmea homozigota para os olhos vermelhos (b^+b^+) é cruzada com um macho com olhos castanhos. Como os olhos marrons são recessivos, a mosca-macho tem de ser homozigota para o alelo dos olhos marrons (*bb*). Todos os descendentes deste cruzamento serão heterozigotos (b^+b) e terão olhos vermelhos:

Dica: Para problemas envolvendo múltiplos *loci*, decomponha o cruzamento em dois cruzamentos separados.

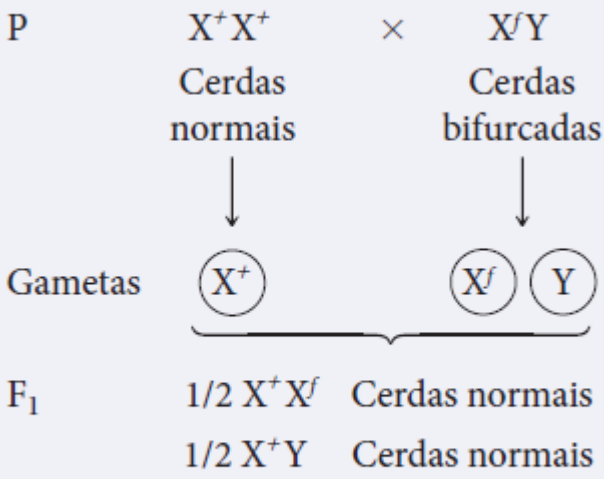


Os descendentes de F_1 são intercruzados para produzir F_2. Sempre que dois organismos heterozigotos para uma característica autossômica recessiva são cruzados, 3/4 dos descendentes terão traço dominante e 1/4 terão traço recessivo; então 3/4 das moscas de F_2 terá olhos vermelhos e 1/4 terá olhos marrons:

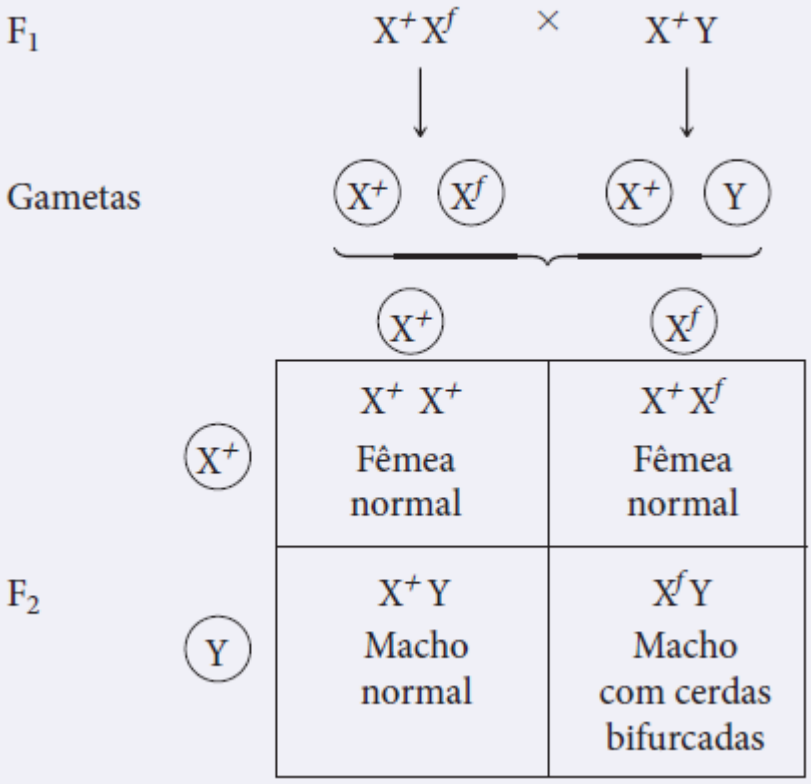


A seguir, vamos trabalhar com os resultados para a característica ligada ao X. Uma mosca-fêmea homozigota para cerdas normais (X^+X^+) é cruzada com uma mosca-macho que tem cerdas bifurcadas (X^fY). A fêmea de F_1 deste cruzamento é heterozigota ($X^+ X^f$) e recebe um cromossomo X com o alelo para cerda normal (X^+) da sua genitora e um cromossomo X com um alelo para cerdas bifurcadas (X^f) de seu genitor. O macho F_1 é hemizigoto (X^+Y) e recebe um cromossomo X com o alelo para cerdas normais (X^+) de sua genitora e um cromossomo Y do seu genitor:

Lembrete: As fêmeas têm dois alelos ligados ao X, mas os machos têm apenas um.



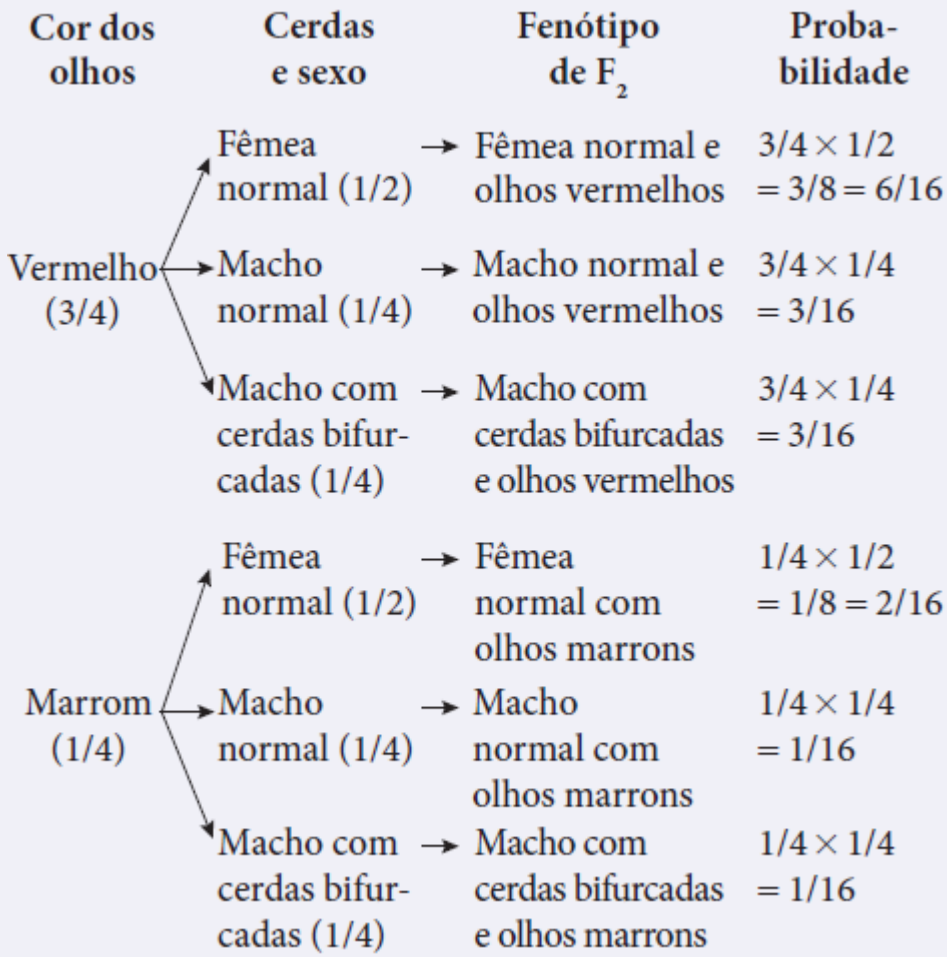
Quando esses descendentes de F_1 são intercruzados, 1/4 de F_2 será de fêmeas com cerdas normais, 1/4 será de machos com cerdas normais e 1/4 será de machos com cerdas bifurcadas:



1/2 fêmea normal, 1/4 macho normal, 1/4 macho com cerdas bifurcadas

Para obter a razão fenotípica em F_2, combinamos esses dois cruzamentos ao usar a regra de multiplicação da probabilidade e o diagrama ramificado:

Lembrete: A regra de multiplicação afirma que a probabilidade de dois eventos independentes ocorrerem juntos é calculada ao multiplicar suas probabilidades independentes.



Dica: O diagrama ramificado é uma forma prática para apresentar todas as combinações diferentes de traços.

Questões de compreensão

Seção 4.1

1. Qual é a diferença fundamental entre machos e fêmeas da maior parte dos organismos?
2. Como os organismos monoicos diferem dos dioicos?
3. Descreva o sistema XX-XO de determinação do sexo. Neste sistema, qual é o sexo heterogamético e qual é o homogamético?

*4. Qual é a diferença da determinação do sexo no sistema XX-XY para o sistema ZZ-ZW?

5. O que é a região pseudoautossômica? Como a herança dos traços codificados pelos genes nesta região é diferente da herança das outras características ligadas ao Y?
6. O que significa a determinação do sexo pelo gene?
7. Qual é a diferença entre a determinação do sexo na *Drosophila* e a determinação de sexo nos seres humanos?
8. Indique o complemento típico de cromossomo sexual encontrado nas células das pessoas com síndrome de Turner, síndrome de Klinefelter e síndrome da insensibilidade aos andrógenos. Qual é o complemento do cromossomo sexual das fêmeas com X triplo?

Seção 4.2

9. Quais características são exibidas por um traço ligado ao X?
10. Explique como o estudo de Bridges sobre a não disjunção na *Drosophila* provou a teoria cromossômica da hereditariedade.
11. Quais características são exibidas por um traço ligado ao Y?

Seção 4.3

12. Explique por que os gatos casco de tartaruga são quase sempre fêmeas e têm distribuição irregular de pelos laranja e preto.
13. O que é um corpúsculo de Barr? Como ele está relacionado com a hipótese de Lyon?

Questões e problemas aplicados

Introdução

*14. Como foi descrito na introdução deste capítulo, os ornitorrincos têm 10 cromossomos sexuais. As fêmeas têm cinco pares de cromossomos X ($X_1X_1X_2X_2X_3X_3X_4X_4X_5X_5$) e os machos têm cinco pares de cromossomos X e Y ($X_1Y_1X_2Y_2X_3Y_3X_4Y_4X_5Y_5$). Cada um dos pares de cromossomo XY se separa de forma independente nos machos durante a meiose? Por que sim ou não?

Seção 4.1

*15. Que fenótipo sexual das moscas-da-fruta tem os seguintes cromossomos?

Cromossomos sexuais	Cromossomos autossômicos
a. XX	Todos normais
b. XY	Todos normais
c. XO	Todos normais
d. XXY	Todos normais
e. XYY	Todos normais
f. XXYY	Todos normais
g. XXX	Todos normais
h. XX	Quatro conjuntos haploides
i. XXX	Quatro conjuntos haploides
j. XXX	Três conjuntos haploides
k. X	Três conjuntos haploides
l. XY	Três conjuntos haploides
m. XX	Três conjuntos haploides

16. Se a não disjunção dos cromossomos sexuais ocorre na meiose no macho na **Figura 4.3**, quais fenótipos sexuais e proporções de descendentes serão produzidos?

*17. Para cada um dos complementos de cromossomo a seguir, qual é o sexo fenotípico de uma pessoa que tem:
   a. XY com gene *SRY* deletado?
   b. XX com uma cópia do gene *SRY* em um cromossomo autossômico?
   c. XO com uma cópia do gene *SRY* em um cromossomo autossômico?
   d. XXY com o gene *SRY* deletado?
   e. XXYY com uma cópia do gene *SRY* deletado?

18. Uma fêmea normal de *Drosophila* produz óvulos anormais com todos (um conjunto completo diploide) os seus cromossomos. Ela cruza com um macho normal de *Drosophila* que produz espermatozoides normais. Qual seria o sexo dos descendentes deste cruzamento?

19. Em algumas salamandras, o sexo de uma fêmea genética pode ser alterado, mudando para um macho funcional, sendo chamadas de machos com sexo invertido. Quando um macho de sexo invertido cruza com uma fêmea normal, aproximadamente 2/3 dos descendentes são fêmeas e 1/3 é macho. Como o sexo é determinado nestas salamandras? Explique os resultados deste cruzamento.

20. Em alguns acarinos, os machos passam os genes para seus netos, mas nunca para seus filhos. Explique.

*21. Nos organismos com o sistema ZZ-ZW de determinação de sexo, a partir de qual das possibilidades a seguir uma fêmea pode herdar seu cromossomo Z?

	Sim	Não
Da avó materna	_____	_____
Do avô materno	_____	_____
Da avó paterna	_____	_____
Do avô paterno	_____	_____

Seção 4.2

*22. Quando Bridges cruzou fêmeas de olhos brancos com machos de olhos vermelhos, ele obteve alguns machos de olhos vermelhos e fêmeas com olhos brancos (ver Figura 4.11). Quais tipos de descendentes seriam produzidos se estes machos de olhos vermelhos e fêmeas de olhos brancos cruzarem entre si?

*23. João tem hemofilia clássica, uma doença recessiva ligada ao X. Ele poderia ter herdado o gene para esta doença a partir das pessoas listadas a seguir?

	Sim	**Não**
A avó materna	_____	_____
O avô materno	_____	_____
A avó paterna	_____	_____
O avô paterno	_____	_____

*24. Na *Drosophila*, o corpo amarelo é produzido a partir de um gene ligado ao X, recessivo para o gene para corpo cinza.
   a. Uma fêmea cinza homozigota é cruzada com um macho amarelo. Os descendentes de F_1 são intercruzados para produzir F_2. Indique os genótipos e fenótipos, junto com as proporções esperadas, dos descendentes de F_1 e F_2.
   b. Uma fêmea amarela é cruzada com um macho cinza. Os descendentes de F_1 são intercruzados para produzir F_2. Indique os genótipos e fenótipos, junto com as proporções esperadas, dos descendentes de F_1 e F_2.
   c. Uma fêmea amarela é cruzada com um macho cinza. As fêmeas de F_1 são retrocruzadas com os machos cinza. Indique os genótipos e fenótipos, junto com as proporções esperadas, dos descendentes de F_2.
   d. Se as moscas de F_2 na parte b se cruzam aleatoriamente, quais são as proporções fenotípicas esperadas das moscas de F_3?

25. A cor do pelo nos gatos é determinada por genes em vários *loci* diferentes. Em um *locus* no cromossomo X, um alelo (X^+) codifica o pelo preto e outro (X^0) codifica o pelo laranja. As fêmeas podem ser pretas ($X^+ X^+$), laranja ($X^0 X^0$) ou uma mistura de laranja e preto chamada de casco de tartaruga (X^+X^0). Os machos são pretos (X^+Y) ou laranja (X^0Y). Bill tem uma fêmea casco de tartaruga chamada de Patches. Uma noite, Patches fugiu da casa de Bill, ficou fora a noite toda e cruzou com um gato de rua. Patches deu à luz os seguintes gatinhos: um macho laranja, um macho preto, duas fêmeas casco de tartaruga e uma fêmea laranja. Indique os genótipos de Patches, seus filhotes e do gato de rua que cruzou com Patches.

*26. O daltonismo vermelho-verde nos seres humanos surge por causa de um gene recessivo ligado ao X. John e Cathy têm visão colorida normal. Após 10 anos de casamento com John, Cathy deu à luz uma menina daltônica. John pediu divórcio, alegando que não era o pai da criança. A alegação de não paternidade de John é justificável? Explique por quê. Se Cathy desse à luz um menino daltônico, a alegação de não paternidade de John seria justificável?

27. O daltonismo vermelho-verde nos seres humanos surge por causa de um gene recessivo ligado ao X. Uma mulher cujo pai é daltônico tem um olho com visão normal e um olho com daltonismo.
   a. Proponha uma explicação para o padrão visual da mulher. Considere que não ocorreram mutações espontâneas.
   b. Seria possível para um homem ter um olho com visão colorida normal e um olho com daltonismo?

*28. Bob tem cromossomos XXY (síndrome de Klinefelter) e é daltônico. Seus pais têm visão colorida normal, mas seu avô materno é daltônico. Considere que a anomalia cromossômica de Bob surgiu a partir de uma não disjunção na meiose. A não disjunção ocorre em qual etapa da divisão meiótica e em qual genitor? Considere que não ocorre *crossing over*. Explique sua resposta.

29. Xg é um antígeno encontrado nas hemácias. Este antígeno é produzido por um alelo ligado ao X (X^a) que é dominante sobre um alelo para a ausência do antígeno (X^-). A herança desses alelos foi estudada em crianças com anomalias cromossômicas para determinar quando ocorria a não disjunção dos cromossomos sexuais. Para cada tipo de cruzamento nas partes de *a* a *d*, indique se ocorreu não disjunção na mãe ou no pai e, se possível, se ela ocorreu na meiose I ou na meiose II (considere que não há *crossing over*).
   a. $X^a Y \times X^- X^- \rightarrow X^a$ (síndrome de Turner).
   b. $X^a Y \times X^a X^- \rightarrow X^-$ (síndrome de Turner).
   c. $X^a Y \times X^- X^- \rightarrow X^a X^- Y$ (síndrome de Klinefelter).
   d. $X^a Y \times X^a X^- \rightarrow X^- X^- Y$ (síndrome de Klinefelter).

30. O Talmude, um livro antigo de leis civis e religiosas judaicas, afirma que, se uma mulher tiver dois filhos que morreram de hemorragia após a circuncisão (remoção do prepúcio do pênis), seus outros filhos não precisam ser circuncisados. É muito provável que o sangramento seja causado pela hemofilia, um distúrbio ligado ao X. Além disso, o Talmude afirma que os filhos de suas irmãs não devem ser circuncisados, enquanto os filhos de seus irmãos devem ser. Esta lei religiosa é coerente com sólidos princípios genéticos? Explique sua resposta.

31. A síndrome craniofacial (CFNS) é um defeito congênito no qual a fusão prematura das suturas cranianas leva a formato anormal do crânio, afastamento dos olhos, fendas nasais e várias outras anomalias do esqueleto. George Feldman e seus colegas examinaram várias famílias nas quais os descendentes tinham CFNS e registraram os resultados apresentados no quadro a seguir (G. J. Feldman. 1997. *Human Molecular Genetics* 6:1937-1941).

			Descendentes			
	Genitores		Normal		CFNS	
Número da família	Pai	Mãe	Homem	Mulher	Homem	Mulher
1	Normal	CFNS	1	0	2	1
5	Normal	CFNS	0	2	1	2
6	Normal	CFNS	0	0	1	2
8	Normal	CFNS	1	1	1	0
10a	CFNS	Normal	3	0	0	2
10b	Normal	CFNS	1	1	2	0
12	CFNS	Normal	0	0	0	1
13a	Normal	CFNS	0	1	2	1
13b	CFNS	Normal	0	0	0	2
7b	CFNS	Normal	0	0	0	2

a. Com base nesses resultados, qual é o modo mais provável de herança para CFNS?
b. Indique os genótipos mais prováveis dos genitores na família 1 e na família 10a.

32. As asas em miniatura (X^m) na *Drosophila* são resultado de um alelo ligado ao X que é recessivo para o alelo de asas longas (X^+). Indique os genótipos dos genitores em cada um dos cruzamentos a seguir.

Genitor macho	Genitora fêmea	Descendentes machos	Descendentes fêmeas
a. Longas	Longas	231 longas, 250 miniaturas	560 longas
b. Miniatura	Longas	610 longas	632 longas
c. Miniatura	Longas	410 longas, 417 miniaturas	412 longas, 415 miniaturas
d. Longas	Miniatura	753 miniaturas	761 longas
e. Longas	Longas	625 longas	630 longas

*33. Nas galinhas, a calvície congênita é causada por um gene recessivo ligado ao Z. Um galo calvo é cruzado com uma galinha normal. Os descendentes desse cruzamento são intercruzados para produzir F_2. Indique os genótipos e fenótipos, junto com as proporções esperadas, de F_1 e F_2.

34. Se as fêmeas F_1 azuis na **Figura 4.13 B** são retrocruzadas com os machos azuis na geração P, quais tipos e proporções de descendentes serão produzidos?

35. O daltonismo das cores vermelho-verde é herdado como um traço recessivo ligado ao X nos seres humanos. A polidactilia (dedos extras nas mãos e pés) é um traço autossômico dominante. Marta tem os dedos das mãos e pés normais, e visão colorida normal. Sua mãe é normal em todos os aspectos, mas seu pai é daltônico e tem polidactilia. Bill é daltônico e tem polidactilia. Sua mãe tem visão colorida normal e dedos das mãos e pés normais. Se Bill e Marta se casarem, quais tipos e proporções de crianças eles terão?

36. Uma mutação na *Drosophila* chamada *queimada* (*s*; do inglês, *singed*) faz com que as cerdas dobrem e se deformem. Uma

mutação chamada *púrpura* (*p*; do inglês *purple*) faz com que os olhos da mosca sejam púrpura em vez do vermelho normal. Moscas homozigotas para cerda queimada e olhos púrpura foram cruzadas com moscas homozigotas para cerdas normais e olhos vermelhos. Os descendentes de F_1 foram intercruzados para produzir F_2 e foram obtidos os resultados a seguir.

Cruzamento 1

P	Macho, cerdas queimadas, olhos púrpura × fêmea, cerdas normais, olhos vermelhos
F_1	420 fêmeas, cerdas normais, olhos vermelhos
	426 machos, cerdas normais, olhos vermelhos
F_2	337 fêmeas, cerdas normais, olhos vermelhos
	113 fêmeas, cerdas normais, olhos púrpura
	168 machos, cerdas normais, olhos vermelhos
	170 machos, cerdas queimadas, olhos vermelhos
	56 machos, cerdas normais, olhos púrpura
	58 machos, cerdas queimadas, olhos púrpura

Cruzamento 2

P	Fêmea, cristas queimadas, olhos púrpura × macho, cerdas normais, olhos vermelhos
F_1	504 fêmeas, cerdas normais, olhos vermelhos
	498 machos, cerdas queimadas, olhos vermelhos
F_2	227 fêmeas, cerdas normais, olhos vermelhos
	223 fêmeas, cerdas queimadas, olhos vermelhos
	225 machos, cerdas normais, olhos vermelhos
	225 machos, cerdas queimadas, olhos vermelhos
	78 fêmeas, cerdas normais, olhos púrpura
	76 fêmeas, cerdas queimadas, olhos púrpura
	74 machos, cerdas normais, olhos púrpura
	72 machos, cerdas queimadas, olhos púrpura

a. Quais são as formas de herança de cerdas queimadas e olhos púrpura? Explique seu raciocínio.
b. Indique os genótipos dos genitores e dos descendentes nas gerações P, F_1 e F_2 do Cruzamento 1 e do Cruzamento 2.

37. Os dois genótipos a seguir são cruzados: *Aa Bb Cc* $X^+ X^r$ × *Aa BB cc* X^+ Y, em que a, b e c representam os alelos para os genes autossômicos e X^+ e X^r representam os alelos ligados ao X em um organismo com a determinação do sexo pelo sistema XX-XY. Qual é a probabilidade de ter o genótipo *aa Bb Cc* X^+X^+ entre os descendentes?

*38. As asas em miniatura na *Drosophila* são produzidas por um gene ligado ao X (X^m), que é recessivo para o alelo de asas longas (X^+). Os olhos marrom-claros são produzidos por um alelo (*s*) que é recessivo para um alelo para olhos vermelhos (s^+).
a. Uma mosca-fêmea com asas miniaturas e olhos marrom-claros é cruzada com um macho com asas normais homozigoto para olhos vermelhos. As moscas de F_1 são intercruzadas para produzir F_2. Indique os fenótipos, assim como suas proporções esperadas, das moscas de F_1 e F_2.
b. Uma mosca-fêmea homozigota para asas normais e olhos marrom-claros é cruzada com um macho com asas em miniatura e homozigoto para olhos vermelhos. As moscas de F_1 são intercruzadas para produzir F_2. Indique os fenótipos, assim como suas proporções esperadas, das moscas de F_1 e F_2.

39. Suponha que um gene recessivo que produz cauda curta nos camundongos esteja localizado na região pseudoautossômica. Um camundongo macho de cauda curta é cruzado com uma fêmea que é homozigota para cauda normal. Os camundongos de F_1 são intercruzados para produzir F_2. Indique os fenótipos, assim como suas proporções esperadas, dos camundongos de F_1 e F_2.

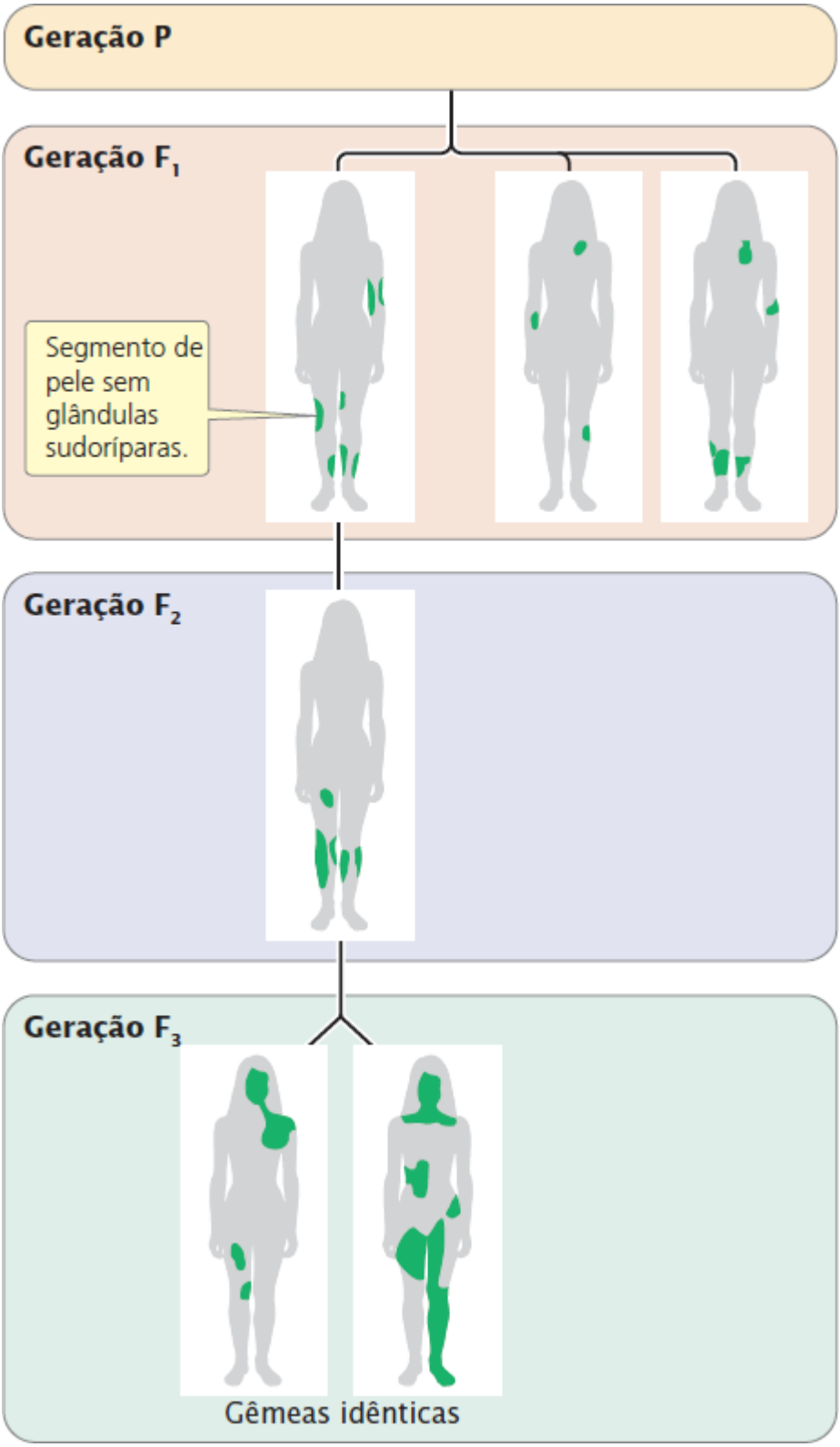
*40. Uma mulher daltônica e um homem com visão normal têm três filhos e seis filhas. Todos os filhos são daltônicos. Cinco das filhas têm visão normal, mas uma é daltônica. A menina daltônica tem 16 anos, é baixa para sua idade e não entrou na puberdade. Explique como esta menina herdou o daltonismo.

Seção 4.3

*41. Quantos corpúsculos de Barr você esperar ver em uma célula humana que tem os seguintes cromossomos?
a. XX

b. XX
c. XO
d. XXY
e. XXXY
f. XXXY
g. XYY
h. XXX
i. XXXX

42. Uma mulher com cromossomos normais tem relações sexuais com um homem que também tem cromossomos normais.
    a. Imagine que, durante a ovogênese, os cromossomos sexuais da mulher sofreram não disjunção na meiose I e os cromossomos do homem se separaram normalmente. Indique todas as possíveis combinações de cromossomos sexuais que as crianças deste casal poderiam herdar e o número de corpúsculos de Barr que você esperaria ver em cada célula de cada criança.
    b. Quais combinações de cromossomo e números de corpúsculos de Barr você esperaria ver se os cromossomos se separassem normalmente na ovogênese, mas a não disjunção dos cromossomos sexuais ocorresse na meiose I da espermatogênese?
43. Quais são o sexo e o genótipo mais prováveis do gato apresentado na **Figura 4.16**?
*44. A displasia ectodérmica anidrótica é um distúrbio recessivo ligado ao X nos seres humanos, caracterizado por dentes pequenos, ausência de glândulas sudoríparas e pouco pelo no corpo. É comum encontrar esse traço nos homens, mas as mulheres heterozigotas portadoras do traço têm segmentos irregulares na pele com poucas ou nenhuma glândula sudorípara (ver a ilustração ao lado).
    a. Explique como as mulheres portadoras heterozigotas de um gene recessivo para a displasia ectodérmica anidrótica têm segmentos irregulares de pele sem glândulas sudoríparas.
    b. Por que a distribuição dos segmentos de pele sem glândula sudorípara é diferente nas mulheres representadas na ilustração, mesmo entre gêmeas idênticas?



(De A. P. Mance e J. Mance, *Genetics: Human Aspects* [Sinauer, 1990], p. 133.)

Questões desafiadoras

Seção 4.2

*45. Um geneticista descobre um camundongo macho com testículos muito aumentados na colônia do seu laboratório. Ele suspeita que esse traço seja o resultado de uma nova mutação que pode ser ligada ao Y ou autossômica dominante. Como ele poderia determinar se o traço é autossômico dominante ou ligado ao Y?

Seção 4.3

46. As mulheres heterozigotas para genes recessivos ligados ao X exibem ocasionalmente uma expressão branda do traço. Entretanto, esta expressão branda dos traços ligados ao X nas fêmeas heterozigotas para os alelos ligados ao X não é

observada na *Drosophila*. O que poderia causar esta diferença na expressão nos genes ligados ao X na mulher e na fêmea da *Drosophila*? (Dica: Na *Drosophila*, a compensação de dosagem é feita ao dobrar a atividade dos genes nos cromossomos X dos machos).

47. Gêmeos idênticos, também chamados de gêmeos monozigóticos, são derivados de um único óvulo fertilizado por um espermatozoide, criando um zigoto que mais tarde se divide em dois (ver *Capítulo 6*). Como os gêmeos idênticos se originam do mesmo zigoto, eles são geneticamente idênticos.

    Caroline Loat e seus colegas examinaram nove medidas de capacidade social, comportamental e cognitiva em 1.000 pares de gêmeos idênticos do sexo masculino e 1.000 pares de gêmeas idênticas do sexo feminino (C. S. Loat *et al.* 2004. *Twin Research* 7:54-61). Eles descobriram que para três das medidas (comportamento pró-social, problemas com os pares e capacidade de se comunicar), os gêmeos do sexo masculino de um par tendiam a ser mais semelhantes nos seus escores do que duas gêmeas de um par. Proponha uma possível explicação para esta observação. O que essa observação pode indicar sobre a localização dos genes que influenciam o comportamento pró-social, problemas com os pares e capacidade de se comunicar?

48. Às vezes, um cromossomo X de camundongo é clivado em dois pedaços e cada fragmento se liga a um cromossomo autossômico diferente. Neste evento, os genes de apenas um dos dois fragmentos sofrem inativação do X. O que essa observação indica sobre o mecanismo de inativação do cromossomo X?

5 Extensões e Modificações dos Princípios Básicos

A estranha genética do caracol canhoto

No início do século 20, o trabalho de Mendel sobre herança nas ervilhas ganhou visibilidade (ver *Capítulo 3*), e vários biólogos começaram a verificar suas conclusões e fizeram cruzamentos com outros organismos. Rapidamente eles confirmaram que os princípios de Mendel se aplicam não apenas a ervilhas, mas também a milho, feijões, galinhas, camundongos, porquinhos-da-índia, seres humanos e muitos outros organismos. Eles, no entanto, também começaram a descobrir exceções, traços cuja herança era mais complexa do que apenas os traços dominantes e recessivos simples que Mendel observou. Uma dessas exceções dizia respeito à espiral da concha do caracol.

O sentido da curvatura do casco do caracol é chamado de quiralidade. A maior parte dos cascos é espiralada no sentido horário, ou para a direita. Esses caracóis são chamados de destros. Alguns caracóis têm cascos espiralados no sentido oposto (anti-horário ou para esquerda), os chamados canhotos (sinistros). A maior parte das espécies de caracóis tem cascos destros ou canhotos; apenas em raros casos os dois tipos coexistem na mesma espécie.

Entre 1920 e 1930, Arthur Boycott, da Universidade de Londres, investigou a genética da curvatura do casco na *Lymnaea peregra*, um caracol comum de lago na Grã-Bretanha. Nessa espécie, a maior parte dos caracóis é destra, mas existem alguns canhotos em algumas populações. Boycott descobriu a partir de naturalistas amadores de um lago perto de Leeds, na Inglaterra, onde poderia encontrar um número anormalmente elevado de caracóis canhotos. Ele conseguiu quatro nesse local e começou a investigar a genética da quiralidade do casco.

A pesquisa de Boycott foi complicada porque esses caracóis são hermafroditas; isso significa que eles podem cruzar com eles mesmos. Se um parceiro adequado estiver disponível, os caracóis também são capazes de cruzar com outro caracol. Boycott descobriu que, se ele isolasse um caracol recémchocado e o criasse sozinho, ele acabaria tendo descendentes, então ele sabia que tinha autocruzado. Mas quando ele colocou dois caracóis juntos e um produziu descendentes, ele não tinha como saber se ele tinha se autocruzado ou cruzado com o outro caracol. A pesquisa de Boycott exigiu vários cultivos de caracóis isolados ou em pares, aumentando o número de descendentes e determinando o sentido do casco para cada um. Para facilitar o trabalho, ele recrutou a ajuda de cientistas amadores. Um de seus assistentes foi o Capitão C. Diver, um amigo que trabalhava como assistente do parlamento Britânico. Como o Parlamento se reunia somente em parte do ano, Diver tinha tempo para auxiliar com a pesquisa. Juntos, Boycott, Diver e outros assistentes realizaram vários cruzamentos, autocruzamentos e entre os caracóis, e organizaram os descendentes em jarros de geleias. Dessa maneira, eles produziram mais de 6.000 descendentes e determinaram o sentido do casco em milhares de caracóis.

Inicialmente, os resultados geraram perplexidade – a curvatura dos cascos não parecia estar em conformidade com os princípios de hereditariedade de Mendel. Assim, eles então descobriram que o lado direito era dominante sobre o lado esquerdo, mas com um truque peculiar: o fenótipo de um caracol era determinado *não* pelo seu genótipo, mas pelo genótipo de sua genitora. Este fenômeno – um fenótipo influenciado pelo genótipo da genitora – é chamado de efeito genético materno. Os efeitos genéticos maternos surgem, em muitos casos, porque a genitora produz uma substância, codificada pelo próprio genótipo, que é depositada no citoplasma do óvulo e influencia o desenvolvimento inicial dos descendentes.

A substância que determina o sentido da curvatura do casco nunca foi isolada. Entretanto, em 2009, Reiko Kuroda e seus colegas demonstraram que o sentido dos cascos nos caracóis *Lymnaea* é determinado pela orientação das células quando o embrião está em um estágio inicial do desenvolvimento, mais especificamente, no estágio de oito células. Ao impulsionar delicadamente as células de embriões de oito células, eles foram capazes de induzir os descendentes cujo genótipo da genitora era destro para se desenvolverem como caracóis canhotos; da mesma maneira, eles induziram os descendentes das genitoras cujo genótipo era canhoto para se desenvolverem como caracóis destros ao impulsionar as células no sentido oposto.

pesquisa de Boycott sobre o sentido do casco nos caracóis demonstrou que nem todas as características são herdadas

A como traços dominante e recessivo simples como o formato e a cor das ervilhas que Mendel descreveu. Essa demonstração não significa que Mendel estava errado, mas que seus princípios não eram por si só suficientes para explicar a herança de todas as características genéticas. Nossa moderna compreensão sobre a genética enriqueceu muito com a descoberta de várias modificações e extensões dos princípios básicos de Mendel, que são o foco deste capítulo.

5.1 Fatores adicionais em um único *locus* podem afetar os resultados dos cruzamentos genéticos

No Capítulo 3, aprendemos que o princípio da segregação e o princípio da segregação independente permitem que possamos prever os resultados dos cruzamentos genéticos. Aqui, examinamos vários fatores adicionais que atuam em *loci* individuais e podem alterar as razões fenotípicas previstas pelos princípios de Mendel.

Tipos de dominância

Uma das contribuições importantes de Mendel para o estudo da hereditariedade é o conceito de dominância, a ideia de que um organismo tem dois diferentes alelos para uma característica, mas o traço codificado por apenas um dos alelos é observado no fenótipo. Com a dominância, o heterozigoto tem o mesmo fenótipo que um dos homozigotos.

Mendel observou a dominância em todos os traços que escolheu estudar, mas ele estava ciente de que nem todas as características exibem dominância. Ele fez alguns cruzamentos sobre o tempo que a ervilha leva para florescer. Por exemplo, quando ele cruzou duas variedades de homozigotos que tinham diferentes períodos de florescência com uma média de 20 dias, o tempo observado para as plantas de F_1 para florescer era intermediário entre os seus dois genitores. Quando o heterozigoto tem um fenótipo intermediário entre os fenótipos dos dois homozigotos, considera-se o traço como *dominância incompleta*.

Dominância completa e dominância incompleta. A dominância pode ser compreendida sobre como o fenótipo do heterozigoto se relaciona com os fenótipos dos homozigotos. No exemplo, apresentado no painel superior da **Figura 5.1**, a cor da flor varia bruscamente de vermelho para branco. Um genótipo homozigoto A^1A^1 produz o pigmento vermelho, que resulta nas flores vermelhas, outro, A^2A^2, não produz pigmentos, resultando em flores brancas. Onde o heterozigoto cai na faixa dos fenótipos determina o tipo de dominância. Se o heterozigoto (A^1A^2) produz a mesma quantidade de pigmento que o homozigoto A^1A^1, resultando em flor vermelha, então o alelo A^1 apresenta **dominância completa** sobre o alelo A^2, ou seja, o vermelho é dominante sobre o branco. Se, por outro lado, o heterozigoto não produz pigmento, e produz flores com a mesma cor do homozigoto A^2A^2 (branco), então o alelo A^2 é completamente dominante, e o branco é dominante sobre o vermelho.

Quando o heterozigoto fica entre os fenótipos dos dois homozigotos, a dominância é incompleta. Com a **dominância incompleta**, o heterozigoto não precisa ser exatamente um intermediário entre os dois homozigotos (ver o painel inferior da Figura 5.1); ele pode ser um tom mais claro de vermelho ou um tom levemente rosado do branco. Desde que o fenótipo do heterozigoto possa ser diferenciado e ficar dentro da faixa de dois homozigotos, a dominância é incompleta.

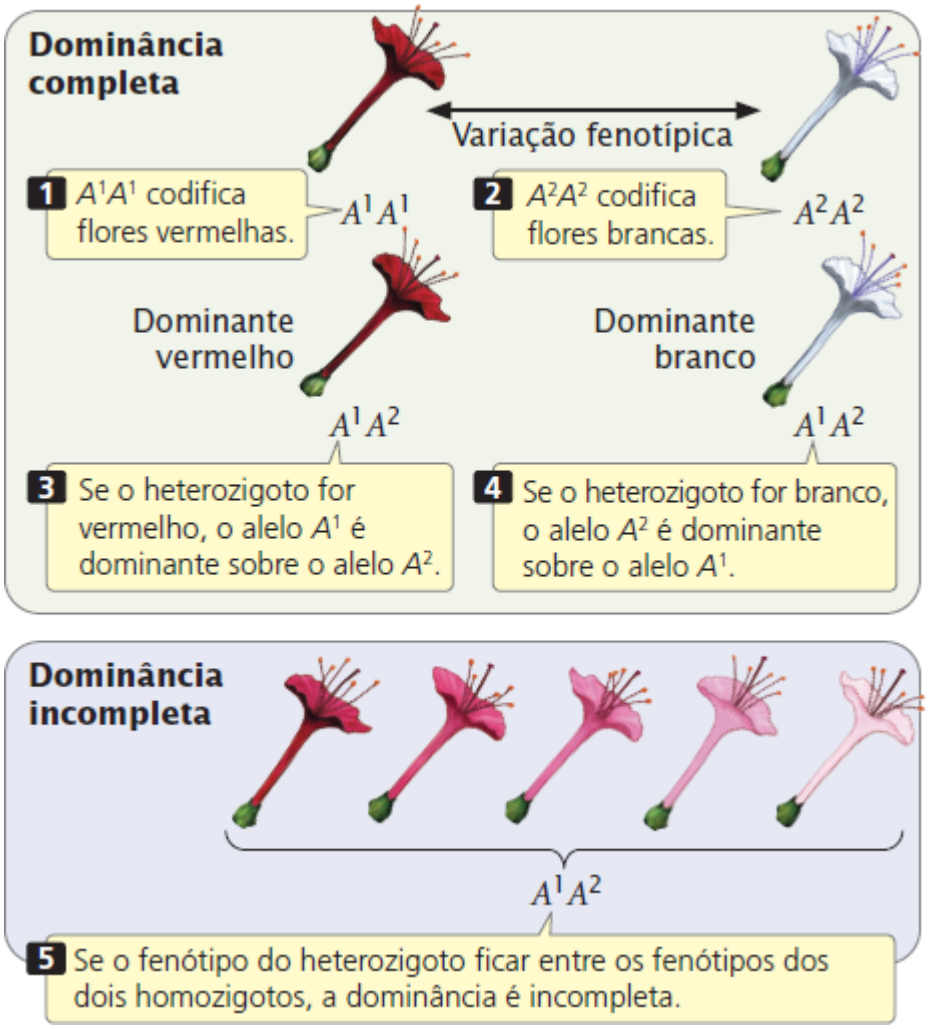


Figura 5.1 O tipo de dominância por um traço depende de como o fenótipo do heterozigoto se relaciona com os fenótipos dos homozigotos.

A dominância incompleta também é exibida na cor da berinjela. Quando uma planta homozigota que produz um fruto púrpura (*PP*) é cruzada com uma planta homozigota que produz um fruto branco (*pp*), todos os heterozigotos (*Pp*) produzem o fruto violeta (**Figura 5.2 A**). Quando os descendentes F_1 são intercruzados, 1/4 de F_2 é púrpura (*PP*), 1/2 é violeta (*Pp*) e 1/4 é branco (*pp*), como demonstrado na **Figura 5.2 B**. Observe que essa razão 1:2:1 é diferente da razão 3:1 que seria observada se a cor da berinjela exibisse dominância completa.

Outro exemplo de dominância incompleta é a cor da pena das galinhas. Um cruzamento entre uma galinha preta homozigota e uma galinha homozigota branca produz galinhas de F_1 que são cinza. Se essas galinhas cinza F_1 são intercruzadas, elas produzem aves de F_2 em uma razão de 1 preto: 2 cinza: 1 branco.

Agora devemos adicionar a razão 1:2:1 a essas razões genotípicas por simples cruzamentos apresentados no Capítulo 3 (ver Quadro 3.5). Uma razão fenotípica 1:2:1 surge nos descendentes de um cruzamento entre dois genitores heterozigotos para uma característica que exibe dominância incompleta (*Aa* × *Aa*). A razão genotípica entre esses descendentes também é 1:2:1. Quando um traço apresenta dominância incompleta, as razões genotípicas e fenotípicas dos descendentes são as *mesmas*, porque cada genótipo tem seu próprio fenótipo. É importante lembrar sobre a dominância que ela afeta a forma como os genes são *expressos* (o fenótipo), mas não a forma como os genes são *herdados*.

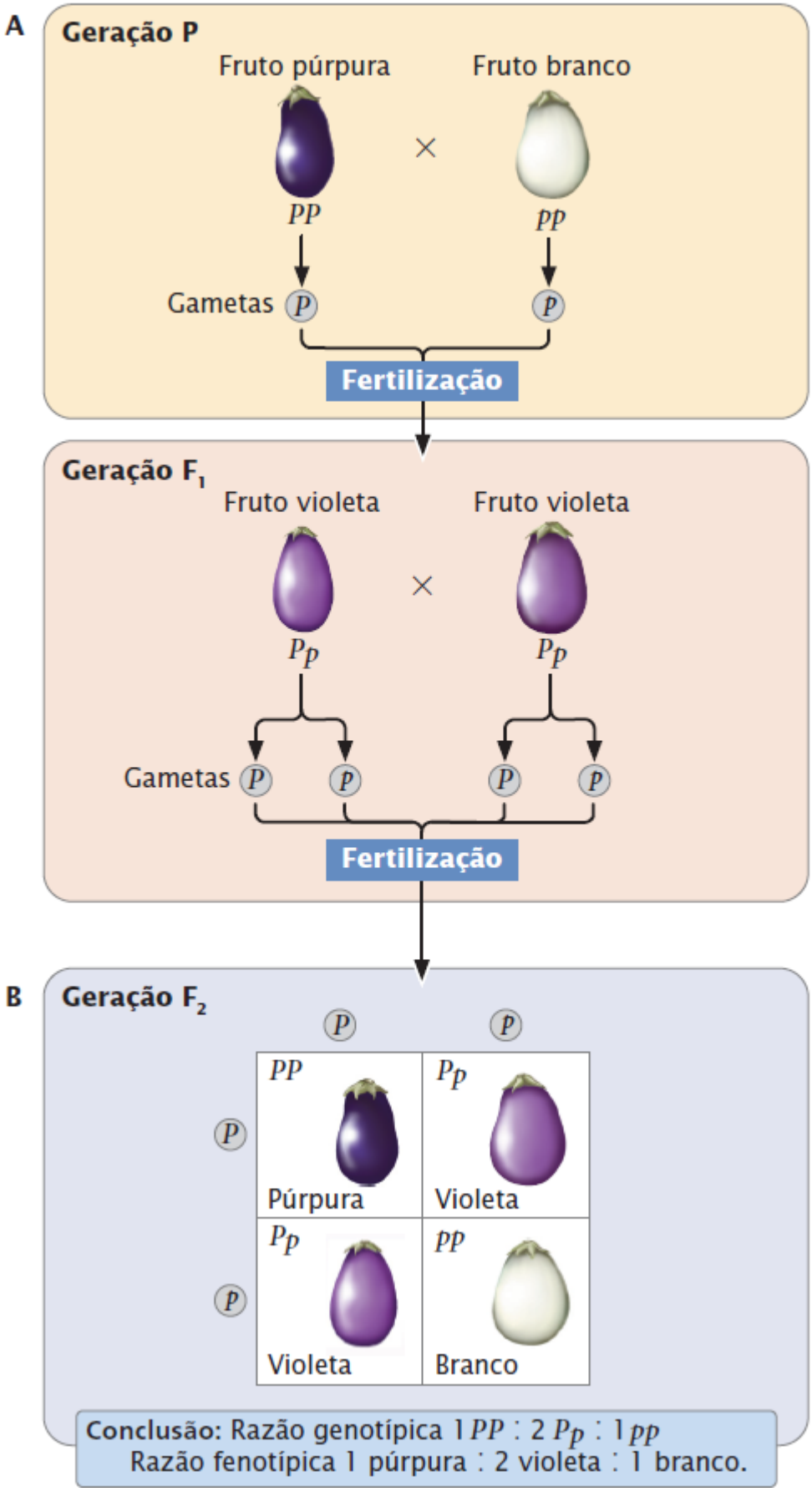


Figura 5.2 A cor da berinjela é herdada como um traço dominante incompleto.

Conceitos

A dominância incompleta é exibida quando o heterozigoto tem um fenótipo intermediário entre os fenótipos de dois homozigotos. Quando um traço exibe dominância incompleta, um cruzamento entre dois

heterozigotos produz uma razão fenotípica 1:2:1 nos descendentes.

Checagem dos conceitos 1

Se uma berinjela de F_1 na Figura 5.2 for usada em um cruzamento-teste, qual a proporção de descendentes deste cruzamento será branca?

a. Todos os descendentes.
b. 1/2.
c. 1/4.
d. 0.

Codominância. Outro tipo de interação entre os alelos é a **codominância**, na qual o fenótipo do heterozigoto não é intermediário entre os fenótipos dos homozigotos, mas o heterozigoto expressa simultaneamente os fenótipos de ambos os homozigotos. Um exemplo de codominância é observado no tipo sanguíneo MN.

O *locus* MN codifica um dos tipos de antígenos nas células vermelhas do sangue. Diferente dos antígenos estranhos aos grupos AB0 e Rh (que também codificam antígenos dessas células), os antígenos estranhos MN não produzem uma reação imunológica forte, portanto, os tipos sanguíneos MN não são levados em consideração na rotina de transfusão de sangue. Existem dois alelos no *locus* MN: o alelo L^M, que codifica o antígeno M; e o alelo L^N, que codifica o antígeno N. Os homozigotos com genótipo L^ML^M expressam o antígeno M nas suas células vermelhas do sangue e têm o tipo sanguíneo M. Os homozigotos com o genótipo L^NL^N expressam o antígeno N e têm o tipo sanguíneo N. Os heterozigotos com genótipo L^ML^N exibem a codominância e expressam os antígenos M e N; seu tipo sanguíneo é MN.

Alguns estudantes podem perguntar por que as flores rosa ilustradas no painel inferior da Figura 5.1 exibem dominância incompleta, ou seja, por que esse resultado não é um exemplo de codominância? As flores exibiriam codominância apenas se o heterozigoto produzisse os pigmentos vermelho e branco, que então se combinam para produzir o fenótipo rosa. Entretanto, em nosso exemplo, o heterozigoto produz apenas o pigmento vermelho. O fenótipo rosa surge porque a quantidade de pigmento produzido pelo heterozigoto é menor que a quantidade produzida pelo homozigoto A^1A^1. Então, neste exemplo, os alelos exibem claramente a dominância incompleta, e não a codominância. As diferenças entre dominância completa, dominância incompleta e codominância estão resumidas no **Quadro 5.1**. Resolva o Problema 13

O nível do fenótipo observado pode afetar a dominância. Os fenótipos podem ser observados com frequência em vários níveis diferentes, incluindo os níveis anatômico, fisiológico e molecular. O tipo de dominância exibido por uma característica depende do nível do fenótipo examinado. Essa dependência é observada na fibrose cística, um distúrbio genético comum nos caucasianos e em geral considerado uma doença recessiva. As pessoas com fibrose cística produzem muito muco espesso e viscoso, que obstrui as vias respiratórias dos pulmões e entope os ductos que saem do pâncreas para o intestino, causando infecções respiratórias e problemas digestivos frequentes. Mesmo com tratamento médico, os pacientes com fibrose cística sofrem com problemas crônicos e risco de morte.

Quadro 5.1 Diferenças entre dominância completa, dominância incompleta e codominância.

Tipo de dominância	Definição
Dominância completa	O fenótipo do heterozigoto é o mesmo fenótipo que um dos homozigotos
Dominância incompleta	O fenótipo do heterozigoto é intermediário (fica dentro da faixa) entre os fenótipos dos dois homozigotos
Codominância	O fenótipo do heterozigoto inclui os fenótipos de ambos homozigotos

O gene responsável pela fibrose cística reside no braço longo do cromossomo 7. Ele codifica uma proteína chamada de *regulador da condutância transmembrana da fibrose cística* (CFTR), que age como um portal na membrana celular e regula o movimento dos íons de cloro para dentro e para fora da célula. As pessoas com fibrose cística têm uma forma mutante, não funcional da CFTR que faz com que o canal permaneça fechado e então os íons de cloro se acumulem na célula. Esse acúmulo provoca a formação de muco espesso e produz os sintomas da doença.

A maior parte das pessoas tem duas cópias do alelo normal para CFTR e produz apenas a proteína CFTR funcional. As

pessoas com fibrose cística têm duas cópias do alelo mutante de CFTR e produzem apenas a proteína CFTR defeituosa. Os heterozigotos, que têm um alelo para CFTR normal e um alelo para CFTR defeituoso, produzem ambas as proteínas CFTR, normal e defeituosa. Desta forma, em nível molecular, os alelos para CFTR normal e defeituoso são codominantes, porque ambos os alelos são expressos no heterozigoto. Entretanto, um alelo produz proteína CFTR funcional o suficiente para permitir o transporte normal do íon cloreto, de modo que o heterozigoto não exibe os efeitos adversos e o alelo do CFTR mutante parece ser recessivo em nível fisiológico. O tipo de dominância expresso por um alelo, como ilustrado neste exemplo, é uma função do aspecto fenotípico do alelo que é observado.

Características da dominância. Várias características importantes da dominância devem ser enfatizadas. Primeiro, a dominância é um resultado de interações gênicas no mesmo *locus*, em outras palavras, a dominância é uma interação *alélica*. Segundo, a dominância não altera a forma na qual os genes são herdados, ela apenas influencia a forma na qual eles são expressos como um fenótipo. A interação alélica que caracteriza a dominância é, portanto, a interação entre os *produtos* dos genes. Finalmente, a dominância está "nos olhos de quem vê", o que significa que a classificação de dominância depende do nível no qual o fenótipo é examinado. Como foi observado para a fibrose cística, um alelo pode exibir codominância em um nível e ser recessivo em outro nível.

Conceitos

A dominância vincula as interações gênicas no mesmo *locus* (genes alélicos) e é um aspecto do fenótipo, não afetando a forma na qual os genes são herdados. O tipo de dominância exibido por uma característica depende do nível do fenótipo examinado.

Checagem dos conceitos 2

Qual é a diferença entre dominância completa, dominância incompleta e codominância?

Penetrância e expressividade

Nos cruzamentos genéticos apresentados até o momento, consideramos apenas as interações dos alelos e consideramos que todos os organismos têm um genótipo particular que expressa o fenótipo esperado. Supomos, por exemplo, que o genótipo *Rr* sempre produz sementes lisas e que o genótipo *rr* sempre produz sementes rugosas. Para algumas características, entretanto, tal hipótese está errada: nem sempre o genótipo produz o fenótipo esperado, um fenômeno chamado de **penetrância incompleta**.

A penetrância incompleta é observada nas pessoas com polidactilia, a condição que produz dedos extras nos pés ou mãos (**Figura 5.3**). Existem várias formas diferentes de polidactilia humana, mas em geral o traço é causado por um alelo dominante. Às vezes, as pessoas têm o alelo para polidactilia (como confirmado pelo fato de que seus filhos herdam a polidactilia), mas têm o número normal de dedos. Nesses casos, o gene para polidactilia não é totalmente penetrante. A **penetrância** é definida como a porcentagem de organismos individuais com um genótipo particular que expressa o fenótipo esperado. Por exemplo, se examinarmos 42 pessoas com um alelo para polidactilia e descobrirmos que apenas 38 destas têm polidactilia, a penetrância seria 38/42 = 0,90 (90%).

Um conceito relacionado é o da **expressividade**, o grau no qual um traço é expresso. Além da penetrância incompleta, a polidactilia exibe expressividade variável. Algumas pessoas com esse quadro têm dedos extras nos pés e nas mãos que são totalmente funcionais, enquanto algumas têm apenas um pequeno pedaço de pele a mais.

A penetrância incompleta e expressividade variável são causadas pelos efeitos de outros genes e fatores ambientais que podem alterar ou suprimir completamente o efeito de um gene particular. Por exemplo, um gene pode codificar uma enzima que produz um fenótipo específico em uma faixa de temperatura limitada. Em temperaturas maiores ou menores, a enzima não funciona e o fenótipo não é expresso; o alelo que codifica esta enzima é, portanto, penetrante apenas dentro de uma faixa de temperatura particular (ver, adiante, *Efeitos ambientais no fenótipo*). Muitas características exibem penetrância incompleta e expressividade variável, assim, a simples presença de um gene não garante sua expressão. **Resolva o Problema 15**

Figura 5.3 A polidactilia humana (dedos extras) exibe penetrância incompleta e expressividade variável. (SPL/Photo Researchers.)

Conceitos

A penetrância é a porcentagem de organismos com um genótipo particular que expressa o fenótipo associado. A expressividade é o grau no qual um traço é expresso. A penetrância incompleta e a expressividade variável resultam na influência de outros genes e fatores ambientais no fenótipo.

Checagem dos conceitos 3

Qual é a diferença entre dominância incompleta e penetrância incompleta?

a. A dominância incompleta se refere aos alelos no mesmo *locus*; a penetrância incompleta se refere a alelos em diferentes *loci*.
b. A dominância incompleta varia de 0 a 50%, a penetrância incompleta varia de 51 a 99%.
c. Na dominância incompleta, o heterozigoto é intermediário entre os homozigotos; na penetrância incompleta, o heterozigoto expressa os fenótipos de ambos os homozigotos.
d. Na dominância incompleta, o heterozigoto é intermediário entre os homozigotos; na penetrância incompleta, alguns indivíduos não expressam o fenótipo esperado.

Alelos letais

Um **alelo letal** provoca a morte no estágio inicial do desenvolvimento – em geral antes do parto – de modo que alguns genótipos podem não aparecer entre os descendentes. Um exemplo de um alelo letal, originalmente descrito por Erwin Baur em 1907, é encontrado na planta boca-de-dragão. A cepa *dourada* nessas plantas tem folhas amarelas. Quando duas plantas com folhas amarelas são cruzadas, 2/3 dos descendentes têm folhas amarelas e 1/3 tem folhas verdes. Quando o verde é cruzado com verde, todos os descendentes têm folhas verdes, entretanto, quando o amarelo é cruzado com verde, 1/2 dos descendentes tem folhas verdes e 1/2 tem folhas amarelas, confirmando que todas as plantas boca-de-dragão de folha amarela são heterozigotas.

Outro exemplo de alelo letal é um que determina o pelo amarelo nos camundongos. Um cruzamento entre dois camundongos amarelos heterozigotos produz uma razão genotípica inicial de 1/4 *YY*, 1/2 *Yy* e 1/4 *yy*, mas os camundongos homozigotos *YY* morriam no início do desenvolvimento e não aparecem entre os descendentes, resultando em uma razão 2:1 de *Yy* (amarelo) a *yy* (não amarelo) nos descendentes (**Figura 5.4**). Uma razão 2:1 é quase sempre produzida por um alelo letal recessivo, então observar esta razão entre os descendentes de um cruzamento entre os indivíduos com o mesmo fenótipo é uma forte pista de que

um dos alelos seja letal. Neste exemplo, como no exemplo das folhas amarelas na planta boca-de-dragão, o alelo letal é recessivo porque ele provoca a morte apenas nos homozigotos. Diferente do seu efeito na *sobrevivência*, o efeito do alelo na *cor* é dominante nos camundongos e nas plantas boca-de-dragão, assim, uma única cópia do alelo no heterozigoto produz a cor amarela. Este exemplo ilustra a questão levantada anteriormente (p. 96) de que o tipo de dominância depende do aspecto do fenótipo examinado.

Muitos alelos letais na natureza são recessivos, mas os alelos letais também podem ser dominantes, de modo que os homozigotos e heterozigotos para esse alelo morrem. Os alelos letais realmente dominantes não podem ser transmitidos, exceto se eles forem expressos após o início da reprodução. Resolva o Problema 17

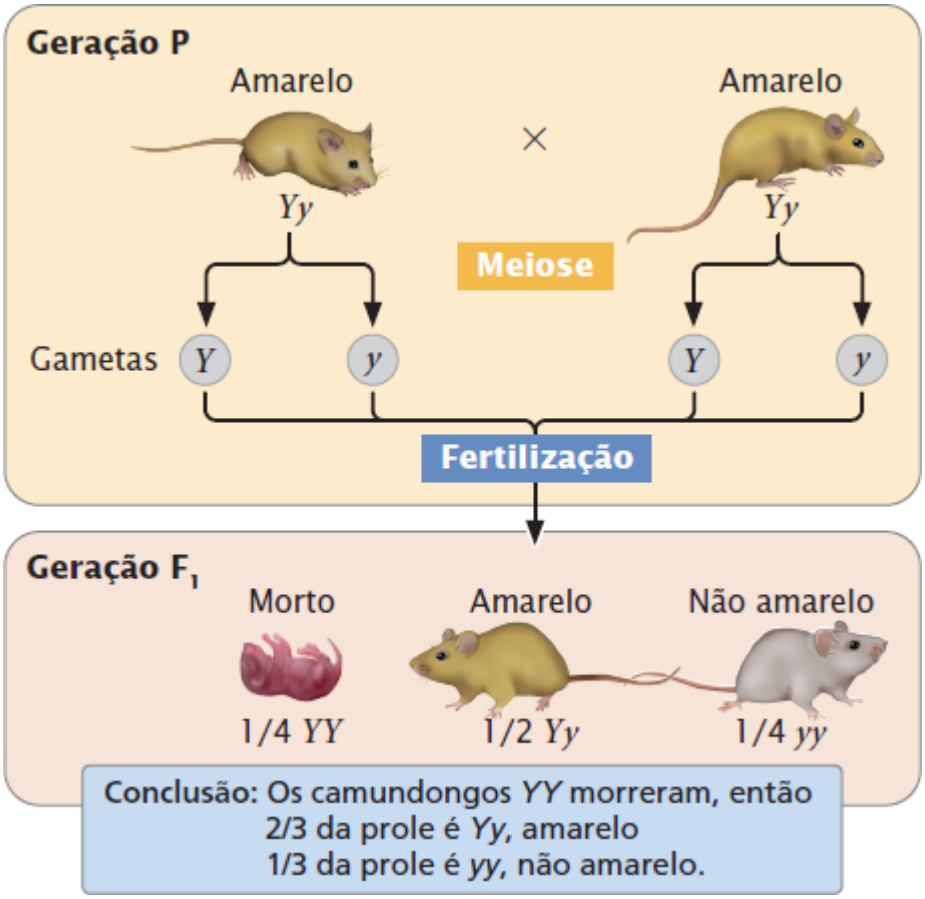


Figura 5.4 A razão 2:1 produziu um cruzamento entre dois camundongos amarelos a partir de um alelo letal.

Conceitos

Um alelo letal provoca a morte, frequentemente no estágio inicial do desenvolvimento, então um ou mais genótipos estão ausentes na prole de um cruzamento. Os alelos letais modificam a razão de descendentes.

Checagem dos conceitos 4

Um cruzamento entre dois milhos-verdes gera 2/3 de descendentes que são verdes e 1/3 dos descendentes que é amarelo. Qual é o genótipo dos descendentes verdes?

a. *WW*.
b. *Ww*.
c. *ww*.
d. *W_* (*WW* e *Ww*).

Alelos múltiplos

A maior parte dos sistemas genéticos que examinamos até aqui tem dois alelos. Nas ervilhas de Mendel, por exemplo, um alelo codifica as sementes lisas e outro codifica as rugosas; nos gatos, um alelo produz o pelo preto e outro, o pelo cinza. Para alguns *loci*, mais de dois alelos estão presentes dentro de um grupo de organismos – o *locus* tem **alelos múltiplos** (também chamados de *série alélica*). Embora possam existir mais de dois alelos dentro de um *grupo* de organismos, o genótipo de cada organismo

diploide ainda terá apenas dois alelos. A herança das características codificadas pelos alelos múltiplos não é diferente da herança das características codificadas por dois alelos, exceto que é possível uma maior variedade de genótipos e fenótipos.

Padrões da pena do pato. Um exemplo de alelos múltiplos está em um *locus* que determina o padrão de pena dos patos *mallard* (pato-real). Um alelo, *M*, produz o padrão selvagem *mallard*. Um segundo alelo, M^R, produz um padrão diferente chamado de *restrito* e um terceiro alelo, m^d, produz um padrão chamado de *escuro*. Nesta série alélica, o restrito é dominante sobre o *mallard* e o escuro e o *mallard* é dominante sobre o escuro: $M^R > M > m^d$. Os seis genótipos possíveis com estes três alelos e seus fenótipos resultantes são:

Genótipo	Fenótipo
M^RM^R	Restrito
M^RM	Restrito
M^Rm^d	Restrito
MM	*Mallard*
Mm^d	*Mallard*
m^dm^d	Escuro

Em geral, o número de genótipos possíveis será $[n(n+1)/2]$, em que n é igual ao número de diferentes alelos em um *locus*. Fazer cruzamentos com alelos múltiplos não é diferente de fazer cruzamentos com dois alelos; o princípio da segregação de Mendel ainda funciona, como é demonstrado no cruzamento entre um pato restrito e um pato *mallard* (**Figura 5.5**).

Resolva o Problema 19

O grupo sanguíneo AB0. Outro sistema de alelos múltiplos está no *locus* para o grupo sanguíneo AB0. Este *locus* determina seu tipo sanguíneo AB0 como o *locus* para MN codifica os antígenos nos eritrócitos do sangue. Os três alelos comuns para o grupo sanguíneo AB0 são: I^A, que codifica o antígeno A, I^B, que codifica o antígeno B e *i*, que não codifica antígeno (0). Podemos representar as relações de dominância entre os alelos AB0 da seguinte maneira: $I^A > i$, $I^B > i$, $I^A = I^B$. Ambos os alelos I^A e I^B são dominantes sobre *i* e codominantes um com o outro. O fenótipo AB ocorre graças à existência de um alelo I^A e um alelo I^B, que resulta na produção dos antígenos A e B nos eritrócitos. Uma pessoa com genótipo *ii* não produz antígeno e tem o tipo sanguíneo 0. Os seis genótipos comuns neste *locus* e seus fenótipos estão demonstrados na **Figura 5.6 A.**

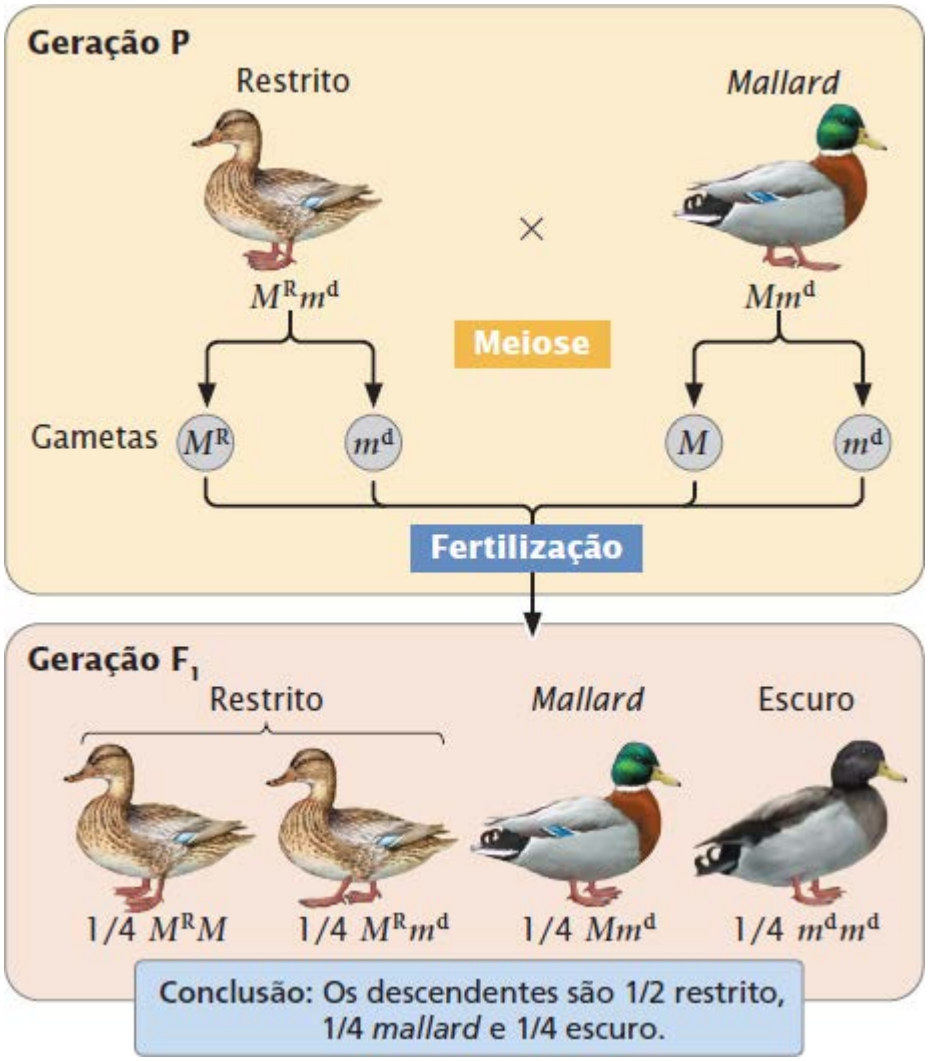


Figura 5.5 O princípio de segregação de Mendel se aplica a cruzamentos com alelos múltiplos. Neste exemplo, três alelos determinam o tipo de plumagem nos patos *mallard*: M^R (restrito) > M (*mallard*) > m^d (escuro).

Os anticorpos são produzidos contra antígenos estranhos (ver Figura 5.6 A). Por exemplo, uma pessoa com tipo sanguíneo A produz anticorpos anti-B, porque o antígeno B é estranho. Uma pessoa com tipo sanguíneo B produz anticorpos anti-A e as pessoas com tipo AB não produzem anticorpos anti-A nem anti-B, porque nenhum destes dois antígenos é estranho. Uma pessoa com tipo sanguíneo 0 não tem os antígenos A ou B; consequentemente, ela produz anticorpos anti-A e anti-B. A existência de anticorpos contra antígenos AB0 estranhos significa que é possível fazer transfusão sanguínea apenas entre pessoas com alguns tipos sanguíneos compatíveis (**Figura 5.6 B**).

A herança dos alelos no *locus* AB0 pode ser ilustrada por meio do processo de paternidade contra o famoso ator Charlie Chaplin. Em 1941, Chaplin conheceu uma jovem atriz chamada Joan Barry, com quem teve um romance. O relacionamento terminou em fevereiro de 1942, mas, 20 meses depois, Barry deu à luz uma menina e afirmou que Chaplin era o pai. E então entrou na justiça pedindo auxílio financeiro. Naquela época, a tipagem sanguínea já era adotada, e os advogados de Chaplin pediram o tipo sanguíneo de Chaplin, Barry e sua filha. Ela era A, sua filha era B e Chaplin era 0. Chaplin poderia ser o pai da criança?

Sua resposta é não. Joan Barry tinha tipo A, que pode ser produzido pelo genótipo $I^A I^A$ ou $I^A i$. Sua filha era B, que pode ser produzido pelo genótipo $I^B I^B$ ou $I^B i$. O bebê não poderia ter herdado o alelo I^B de Barry (ela não pode ter o alelo I^B, já que era A); portanto o bebê necessariamente herdou o alelo i de sua mãe. Barry deve ter o genótipo $I^A i$, e o bebê deve ter o genótipo $I^B i$. Como a menina herdou seu alelo i de Barry, ela necessariamente herdou o alelo I^B de seu pai. Como o tipo 0 produz apenas o genótipo ii, Chaplin não poderia ser o pai da filha de Barry. Embora os tipos sanguíneos possam ser usados para excluir a possibilidade de paternidade (como neste caso), eles não podem provar que uma pessoa é pai ou mãe de uma criança, porque muitas pessoas diferentes têm o mesmo tipo sanguíneo.

No curso do julgamento para esclarecer a paternidade contra Chaplin, três patologistas testemunharam que era geneticamente impossível que ele fosse o pai da criança. Contudo, o júri determinou que Chaplin era o pai e ordenou que ele pagasse o auxílio à criança e as custas do processo de Barry. **Resolva o Problema 24**

Heterozigotos compostos. Diferentes alelos dão origem ao mesmo fenótipo. Por exemplo, a fibrose cística (ver p. 95)

surge a partir de defeitos nos alelos do *locus CFTR*, que codifica uma proteína responsável pelo movimento dos íons de cloreto para dentro e para fora da célula. Foram descobertos cerca de 1.000 alelos diferentes para o *locus* de *CFTR* pelo mundo que podem provocar a fibrose cística. Como esta doença é recessiva autossômica, a pessoa normalmente deve herdar dois alelos *CFTR* defeituosos para ter a fibrose cística. Em algumas pessoas, estes dois alelos defeituosos são idênticos, indicando que a pessoa é homozigota. Outras pessoas com fibrose cística são heterozigotas, com dois diferentes alelos defeituosos. Uma pessoa com dois alelos diferentes em um *locus* que resulta em fenótipo recessivo é chamada de **heterozigoto composto**.

A

Fenótipo (tipo sanguíneo)	Genótipo	Tipo de antígeno	Anticorpos produzidos pelo corpo	Reação dos receptores ao sangue do doador: A (anticorpos anti-B)	B (anticorpos anti-A)	AB (sem anticorpos)	0 (anticorpos anti-A e anti-B)
A	I^AI^A ou I^Ai	A	Anti-B				
B	I^BI^B ou I^Bi	B	Anti-A				
AB	I^AI^B	A e B	Nenhum				
0	ii	Nenhum	Anti-A e anti-B				

B

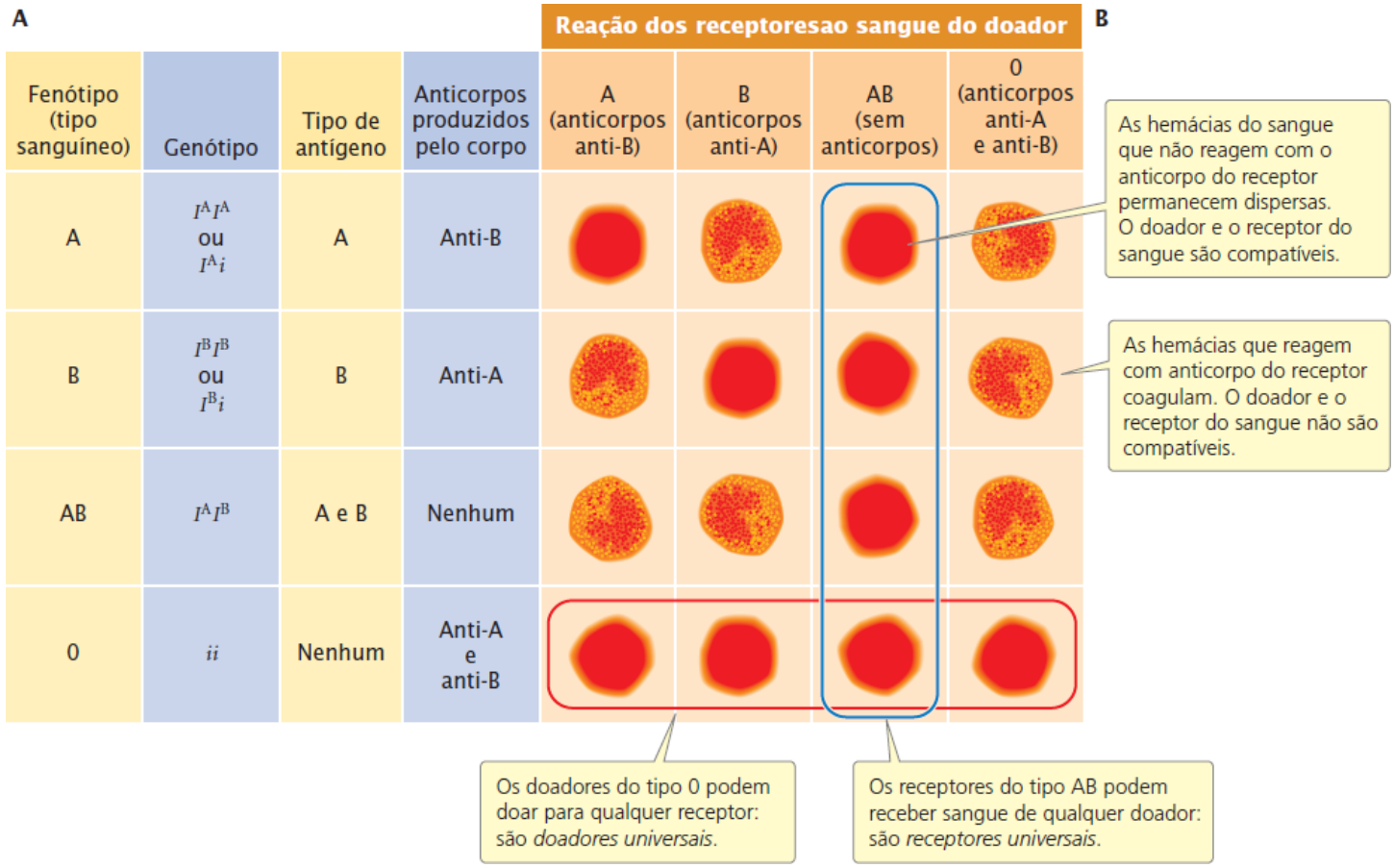


Figura 5.6 Tipos sanguíneos AB0 e possíveis transfusões de sangue.

Conceitos

Mais de dois alelos (alelos múltiplos) podem ser encontrados em um grupo de organismos individuais, embora cada organismo diploide ainda tenha dois alelos naquele *locus*. Um heterozigoto composto tem dois diferentes alelos que resultam em um fenótipo recessivo.

Checagem dos conceitos 5

Quantos genótipos são possíveis em um *locus* com cinco alelos?

a. 30.
b. 27.
c. 15.
d. 5.

5.2 A interação gênica ocorre quando os genes em múltiplos *loci* determinam um único fenótipo

Nos cruzamentos di-híbridos que examinamos no Capítulo 3, cada *locus* tinha um efeito independente sobre o fenótipo. Quando Mendel cruzou uma planta homozigota lisa e amarela (*RR YY*) com uma planta homozigota rugosa e verde (*rr yy*) e então autofertilizou F_1, ele obteve os descendentes F_2 nas seguintes proporções:

9/16 *R_ Y_*	Lisa, amarela
3/16 *R_ yy*	Lisa, verde
3/16 *rr Y_*	Rugosa, amarela
1/16 *rr yy*	Rugosa, verde

Neste exemplo, os genes mostraram dois tipos de independência. Primeiro, os genes em cada *locus* eram independentes na sua *segregação* durante a meiose, o que produziu uma razão 9:3:3:1 de fenótipos nos descendentes, de acordo com o princípio da segregação independente de Mendel. Segundo, os genes são independentes na sua *expressão fenotípica*: os alelos *R* e *r* afetam apenas o formato da semente e não têm influência sobre a cor da semente, os alelos *Y* e *y* afetam apenas a cor e não têm influência sobre o formato da semente.

Frequentemente, os genes exibem a segregação independente, mas não atuam de forma independente na sua expressão fenotípica, pelo contrário, os efeitos dos genes em um *locus* dependem da existência de genes em outros *loci*, interação chamada de **interação gênica**. Com ela, os produtos dos genes em diferentes *loci* se combinam para produzir novos fenótipos que não são previsíveis a partir dos efeitos de um único *locus* sozinho. Em nosso estudo sobre a interação gênica, vamos nos concentrar principalmente na interação entre os efeitos de genes em dois *loci*, embora as interações gênicas em três, quatro ou mais *loci* sejam comuns.

Conceitos

Na interação gênica, genes em diferentes *loci* contribuem para a determinação de uma única característica fenotípica.

Checagem dos conceitos 6

Qual é a diferença entre interação gênica e dominância entre alelos?

A interação gênica e a produção de novos fenótipos

Vamos examinar primeiro a interação gênica na qual os genes em dois *loci* interagem para produzir uma única característica. A cor do fruto do pimentão *Capsicum annuum* é determinada dessa maneira. Alguns tipos de pimentões produzem frutos em uma das quatro cores: vermelha, pêssego, laranja (às vezes chamada de amarela) e creme (ou branca). Se uma planta homozigota com pimentões vermelhos for cruzada com uma planta homozigota com pimentões creme, todas as plantas de F_1 terão pimentões vermelhos (**Figura 5.7 A**). Quando as plantas de F_1 são cruzadas entre si, as plantas de F_2 surgem em uma razão 9 vermelhas: 3 pêssego: 3 laranja: 1 creme (**Figura 5.7 B**). Esta razão di-híbrida (ver *Capítulo 3*) é produzida por um cruzamento entre duas plantas que são heterozigotas para dois *loci* ($Y^+y\ C^+c \times Y^+y\ C^+c$). Neste exemplo, o *locus Y* e o *locus C* interagem para produzir um único fenótipo – a cor do pimentão:

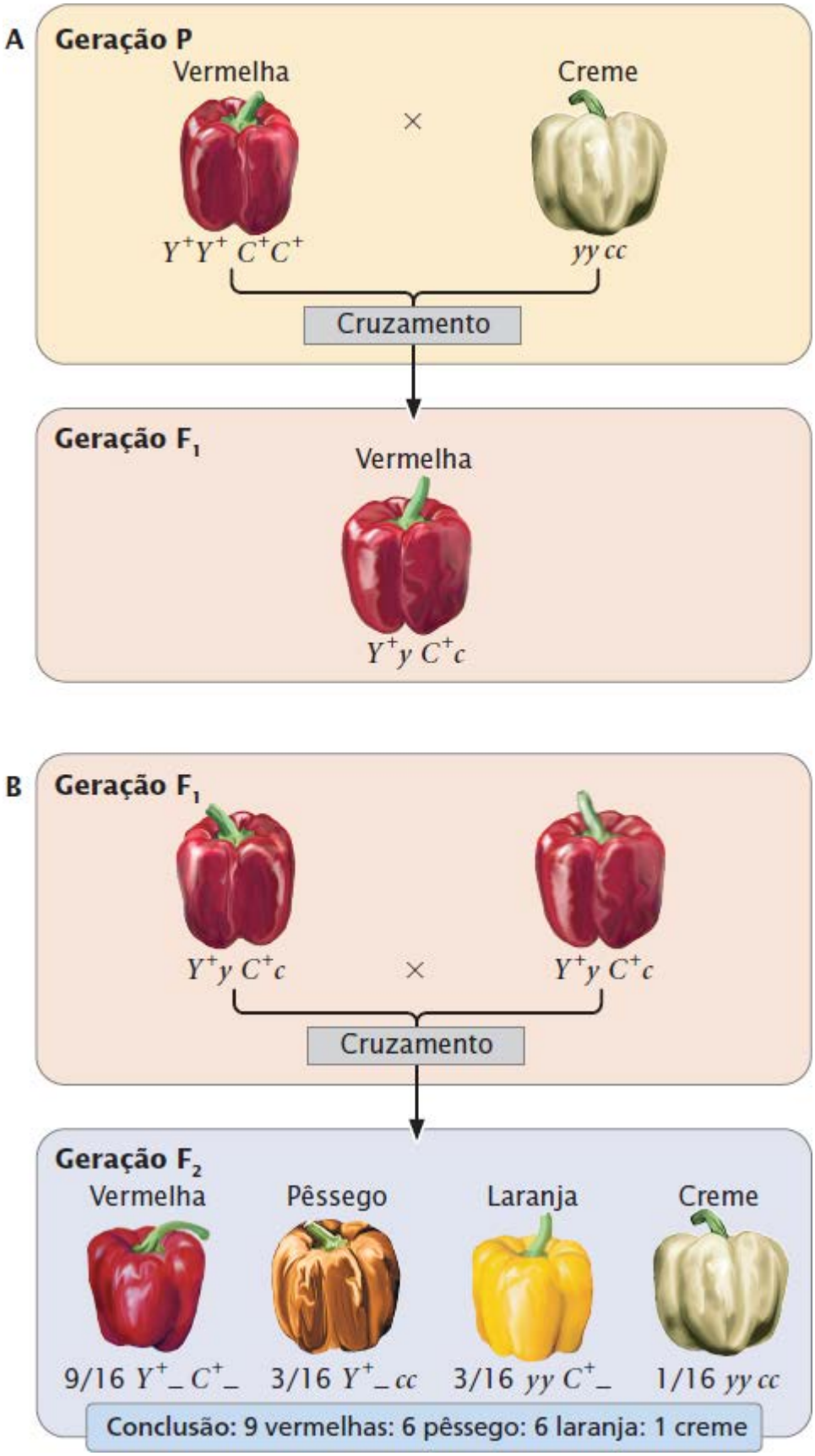


Figura 5.7 A interação gênica em dois *loci* determina uma única característica, a cor da fruta, no caso do pimentão *Capsicum annuum*.

Genótipo	Fenótipo
$Y^+\_\ C^+\_$	Vermelha
$Y^+\_\ cc$	Pêssego
$yy\ C^+\_$	Laranja
$yy\ cc$	Creme

A cor nos pimentões de *Capsicum annuum* é o resultado das concentrações relativas dos carotenoides vermelho e amarelo, compostos sintetizados em uma via bioquímica complexa. O *locus Y* codifica uma enzima (a primeira etapa na via) e o *locus C* codifica uma enzima diferente (a última etapa na via). Quando diferentes *loci* influenciam diferentes etapas em uma via bioquímica comum, a interação gênica surge porque o produto de uma enzima afeta o substrato de outra.

Para ilustrar como as regras de hereditariedade de Mendel podem ser usadas para compreender a herança das características determinadas pela interação gênica, vamos considerar um cruzamento-teste entre as plantas F_1 a partir do cruzamento na Figura 5.7 ($Y^+y\ C^+c$) e uma planta com pimentões creme (*yy cc*). Como foi destacado no Capítulo 3, para os *loci* independentes, podemos trabalhar este cruzamento ao dividi-lo em dois cruzamentos simples. No primeiro *locus*, o heterozigoto Y^+y é cruzado com o homozigoto *yy*, produzindo uma razão de 1/2 Y^+y e 1/2 *yy* de descendentes. Da mesma forma, no segundo *locus*, o genótipo heterozigoto C^+c é cruzado com o homozigoto com genótipo *cc*, gerando uma razão de 1/2 C^+c e 1/2 *cc* de descendentes. De acordo com o princípio de segregação independente de Mendel, essas razões de um *locus* podem ser combinadas ao usar a regra de multiplicação: a probabilidade de obter o genótipo $Y^+y\ C^+c$ é a probabilidade de Y^+y (1/2) multiplicada pela probabilidade de C^+c (1/2) ou 1/2 × 1/2 = 1/4. A probabilidade de cada genótipo dos descendentes oriundos do cruzamento-teste é:

Genótipo dos descendentes	**Probabilidade em cada *locus***		**Probabilidade global**	**Fenótipo**
$Y^+y\ C^+c$	1/2 × 1/2	=	1/4	Pimentões vermelhos
Y^+y *cc*	1/2 × 1/2	=	1/4	Pimentões pêssego
yy C^+c	1/2 × 1/2	=	1/4	Pimentões laranja
yy cc	1/2 × 1/2	=	1/4	Pimentões creme

Quando você lida com problemas de interação gênica, é especialmente importante determinar as probabilidades dos genótipos de único *locus* e multiplicar as probabilidades dos *genótipos*, e não dos fenótipos, porque estes não podem ser determinados sem considerar os efeitos dos genótipos em todos os *loci* que contribuem. **Resolva o Problema 25**

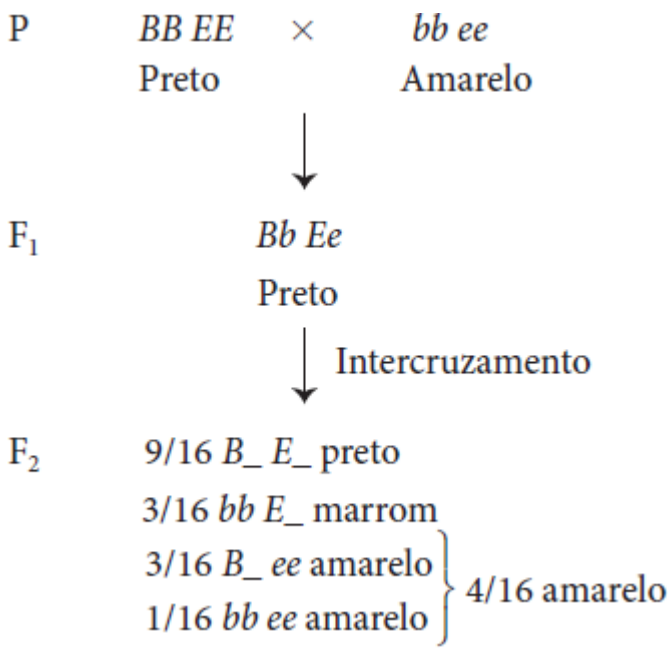
Interação gênica com epistasia

Em alguns casos, o efeito da interação gênica é que um gene mascara (esconde) o efeito de outro gene em um *locus* diferente, um fenômeno conhecido como **epistasia**. Nos exemplos de interação gênica que examinamos, os genes em diferentes *loci* interagem para determinar um único fenótipo. Nesses exemplos, um gene não *mascara* o efeito de outro gene em outro *locus*, o que significa que não há epistasia. Este efeito é semelhante à dominância, exceto que a dominância impõe o mascaramento dos genes no *mesmo locus* (genes alélicos). Na epistasia, o gene que mascara o efeito é chamado de **gene epistático**, o gene cujo efeito é mascarado é um **gene hipostático**. Os genes epistáticos podem ser recessivos ou dominantes em seus efeitos.

Epistasia recessiva. A epistasia recessiva é observada nos genes que determinam a cor do pelo nos labradores retriever. Esses cães podem ser pretos, marrons (frequentemente chamados de chocolate) ou amarelos, e suas diferentes cores são determinadas por interações entre os genes em dois *loci* (embora vários outros *loci* também possam ajudar a determinar a cor do pelo, ver pp. 106-107). Um *locus* determina o tipo de pigmento produzido pelas células da pele: um alelo dominante *B* codifica o pigmento preto, enquanto um alelo recessivo *b* codifica o pigmento marrom. Alelos em um segundo *locus* afetam o *depósito* do pigmento na haste do pelo, o alelo dominante *E* possibilita que o pigmento escuro (preto ou marrom) se deposite, enquanto o alelo recessivo *e* evita o depósito do pigmento escuro, fazendo com que o pelo fique amarelo. A existência do genótipo *ee* no segundo *locus*, portanto, mascara a expressão dos alelos para preto e marrom no primeiro *locus*. Os genótipos que determinam a cor do pelo e seus fenótipos são:

Genótipo	**Fenótipo**
*B_ E_*	Preto
*bb E_*	Marrom
B_ ee	Amarelo
bb ee	Amarelo

Se cruzarmos um labrador preto homozigoto para os alelos dominantes (*BB EE*) com um labrador amarelo homozigoto para os alelos recessivos (*bb ee*) e, então, cruzarmos F_1, obteremos os descendentes em F_2 na razão 9:3:4:



Observe que os cães amarelos podem carrear alelos para o pigmento preto ou marrom, mas eles não expressam sua cor.

Nesse exemplo de interação gênica, o alelo *e* é epistático a *B* e *b*, porque ele mascara a expressão dos alelos para os pigmentos preto e marrom, e os alelos *B* e *b* são hipostáticos a *e*. Nesse caso, *e* é um alelo epistático recessivo, porque duas cópias de *e* devem estar presentes para mascarar a expressão dos pigmentos preto e marrom. Resolva o Problema 29

Outro exemplo de um gene epistático recessivo é o gene que determina o fenótipo Bombaim, que mascara a expressão dos alelos no *locus* AB0. Como já mencionado neste capítulo, os alelos no *locus* AB0 codificam antígenos eritrocitários, os antígenos são pequenas cadeias de carboidratos incrustados nas membranas eritrocitárias. A diferença entre os antígenos A e B é uma função das diferenças químicas no açúcar terminal da cadeia. Os alelos I^A e I^B codificam enzimas diferentes, que adicionam açúcares indicados como A ou B às extremidades das cadeias de carboidrato (**Figura 5.8**). O substrato comum no qual essas enzimas atuam é uma molécula chamada H. A enzima codificada pelo alelo *i* aparentemente não adiciona açúcar a H ou nenhuma enzima funcional está especificada.

Na maior parte das pessoas, um alelo dominante (*H*) no *locus* H codifica uma enzima que produz H, mas as pessoas com o fenótipo Bombaim são homozigotas para uma mutação recessiva (*h*) que codifica uma enzima defeituosa. A enzima defeituosa é incapaz de produzir H e, na ausência deste, não são sintetizados os antígenos do sistema AB0. Assim, a expressão dos alelos no *locus* AB0 depende do genótipo no *locus* H.

Genótipo	H presente	Fenótipo AB0
*H_* I^AI^A, *H_* I^Ai	Sim	A
*H_* I^BI^B, *H_* I^Bi	Sim	B
*H_* I^AI^B	Sim	AB
H_ ii	Sim	0
hh I^AI^A, *hh* I^Ai, *hh* I^BI^B, *hh* I^Bi, *hh* I^AI^B, e *hh ii*	Não	0

Neste exemplo, os alelos no *locus* AB0 são hipostáticos para o alelo recessivo *h*.

O fenótipo Bombaim nos dá uma boa oportunidade para ver como a epistasia surge quando os genes afetam uma série de etapas em uma via bioquímica. Os antígenos AB0 são produzidos em uma via bioquímica de várias etapas (ver Figura 5.8), que dependem das enzimas para produzir H e nenhuma outra enzima que converte H no antígeno A ou B. Observe que o tipo sanguíneo 0 pode surgir em uma das duas formas: (1) a partir da falha para adicionar um açúcar terminal ao composto H (genótipo *H_ ii*) ou (2) a partir da falha para produzir o composto H (genótipo *hh*__). Muitos casos de epistasia surgem dessa forma. Um gene (como *h*) que tem um efeito em um estágio inicial em uma via bioquímica será epistático para os genes (como I^A ou I^B) que afetam as etapas subsequentes, porque os efeitos dos genes em uma etapa posterior dependem do produto da reação inicial.

Epistasia dominante. Na epistasia *recessiva*, a existência de dois alelos recessivos (o genótipo homozigoto) inibe a expressão de um alelo em um diferente *locus*. Entretanto, na epistasia *dominante*, apenas uma única cópia de um alelo é necessária para inibir a expressão do alelo em um *locus* diferente.

A epistasia dominante é observada na interação de dois *loci* que determinam a cor da fruta na abóbora-moranga, que é comumente encontrada em uma de três cores: amarela, branca ou verde. Quando uma planta homozigota que produz abóbora

branca é cruzada com uma planta que produz abóbora verde e as plantas de F_1 são cruzadas entre si, são obtidos os seguintes resultados:

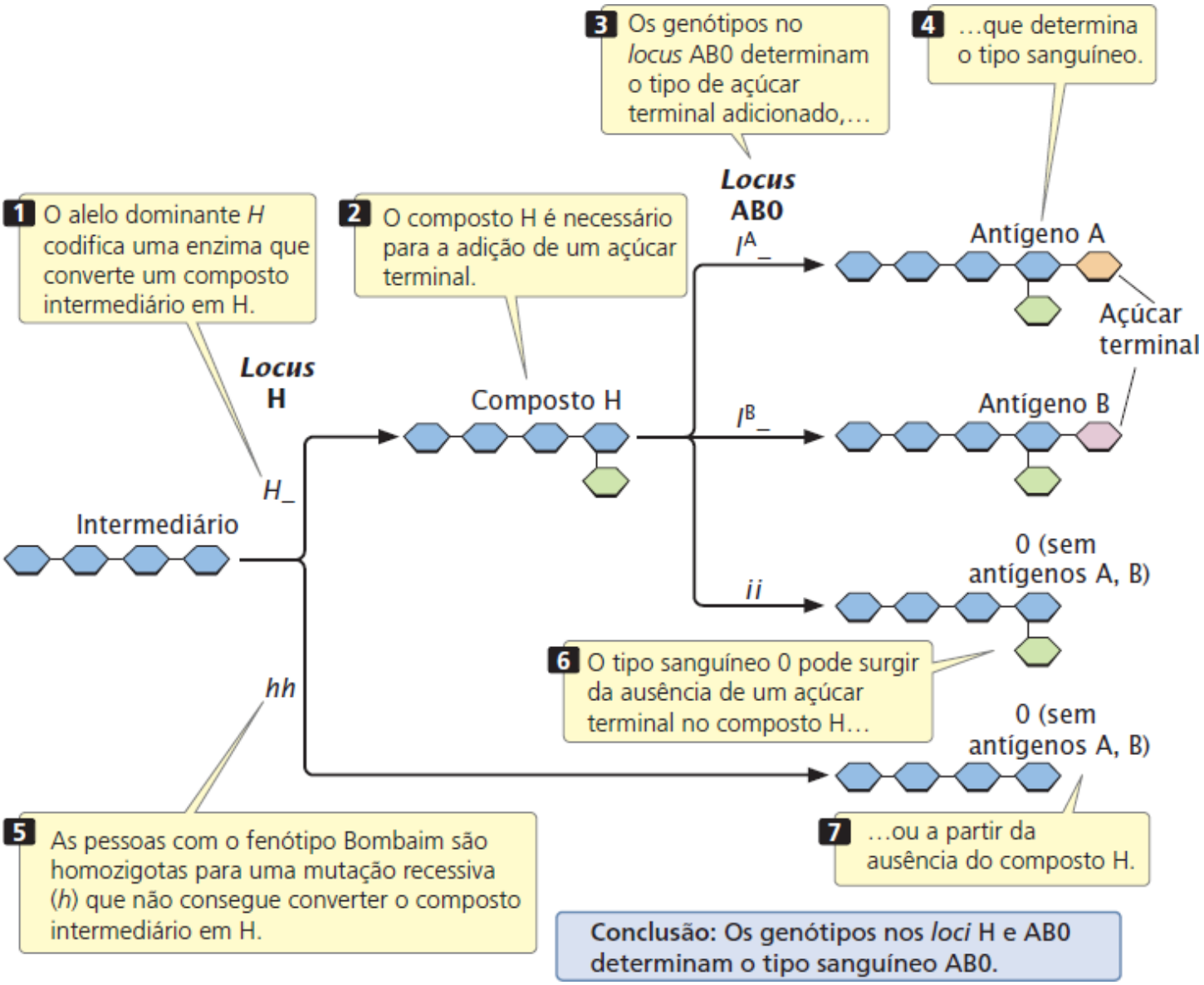


Figura 5.8 A expressão dos antígenos AB0 depende dos alelos no *locus* H. O *locus* H codifica um precursor dos antígenos chamado de composto H. Os alelos no *locus* AB0 determinam quais tipos de açúcares terminais são adicionados ao composto H.

P Plantas com abóbora branca × Plantas com abóbora verde

↓

F_1 Plantas com abóbora branca

↓ Intercruzamento

F_2 12/16 Plantas com abóbora branca

3/16 Plantas com abóbora amarela

1/16 Plantas com abóbora verde

Como a interação gênica pode explicar estes resultados?

Em F_2, 12/16 ou 3/4 das plantas produzem a abóbora branca e 3/16 + 1/16 = 4/16 = 1/4 das plantas produz abóbora com cor. Este desfecho é a familiar razão 3:1 produzida por um cruzamento entre dois heterozigotos, o que sugere que um alelo dominante em um *locus* inibe a produção de pigmento, resultando nos descendentes com fruto branco. Se usarmos o símbolo *W* para representar o alelo dominante que inibe a produção do pigmento, então o genótipo *W_* inibe a produção do pigmento e produz a abóbora branca, enquanto *ww* produz pigmento e consequentemente, a abóbora colorida.

Entre essas plantas *ww* de F_2 com fruto pigmentado, observamos 3/16 de amarelos e 1/16 de verdes (uma razão 3:1). Neste desfecho, um segundo *locus* determina o tipo de pigmento produzido na abóbora, com o amarelo (*Y_*) dominante sobre o verde (*yy*). Esse *locus* é expresso apenas nas plantas *ww*, que não têm o alelo *W* dominante inibitório. Podemos indicar o genótipo *ww Y_* para plantas que produzem abóbora amarela e o genótipo *ww yy* para plantas que produzem abóbora verde. Os genótipos e seus fenótipos associados são:

*W_ Y_*	Abóbora branca
W_ yy	Abóbora branca
*ww Y_*	Abóbora amarela
ww yy	Abóbora verde

O alelo *W* é epistático para *Y* e *y*: ele mascara a expressão desses genes que produzem pigmentos. O alelo *W* é epistático dominante porque, ao contrário de *e* na cor do pelo do labrador retriever e *h* no fenótipo Bombaim, uma única cópia do alelo é suficiente para inibir a produção de pigmento.

É mais provável que o pigmento amarelo na abóbora seja produzido em uma via bioquímica de duas etapas (**Figura 5.9**). Um composto A incolor (branco) (indicado por A na Figura 5.9) é convertido pela enzima I ao composto B verde, que é então convertido no composto C pela enzima II. O composto C é o pigmento amarelo na fruta. As plantas com o genótipo *ww* produzem a enzima I e podem ser verdes ou amarelas, dependendo se contêm a enzima II. Quando o alelo *Y* é encontrado em um segundo *locus*, a enzima II é produzida e o composto B é convertido no composto C, produzindo uma fruta amarela. Quando existem duas cópias do alelo *y*, que não codifica uma forma funcional da enzima II, a abóbora permanece verde. A existência de *W* no primeiro *locus* inibe a conversão do composto A no composto B; as plantas com genótipo *W_* não produzem o composto B e seus frutos permanecem brancos, independentemente de quais alelos estejam presentes no segundo *locus*.

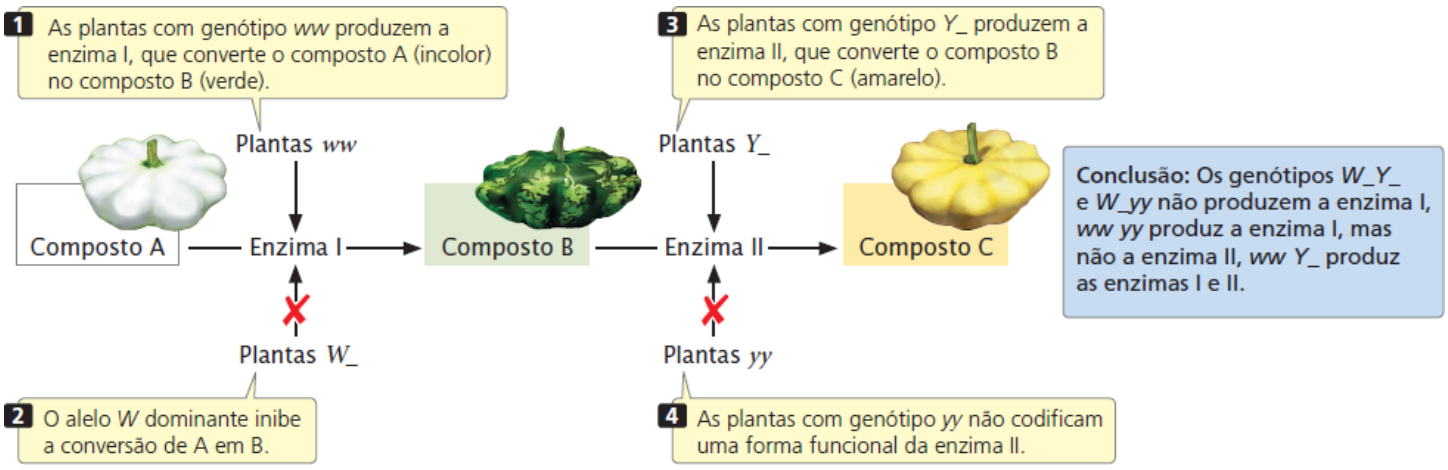


Figura 5.9 **O pigmento amarelo na abóbora é produzido em uma via bioquímica de duas etapas.**

Epistasia recessiva duplicada. Finalmente vamos considerar a epistasia recessiva duplicada, na qual dois alelos recessivos em um dos dois *loci* são capazes de suprimir um fenótipo. Esse tipo de epistasia é demonstrado pelo albinismo nos caracóis.

O albinismo é a ausência de pigmento e é um traço genético comum em muitas plantas e animais. O pigmento é quase sempre produzido por uma via bioquímica de várias etapas; assim, o albinismo exige interação gênica. Robert T. Dillon e Amy R. Wethington descobriram que o albinismo no caracol comum de água doce *Physa heterostropha* pode surgir a partir da existência de dois alelos recessivos em cada um de dois *loci* diferentes. Caracóis inseminados foram coletados de uma população natural e colocados em copos de água, onde depositaram seus ovos. Alguns ovos geraram caracóis albinos. Quando dois caracóis albinos foram cruzados, todos os descendentes F_1 eram pigmentados. Quando os caracóis de F_1 foram intercruzados, o F_2 resultante tinha 9/16 caracóis pigmentados e 7/16 caracóis albinos. Como esta razão 9:7 surgiu?

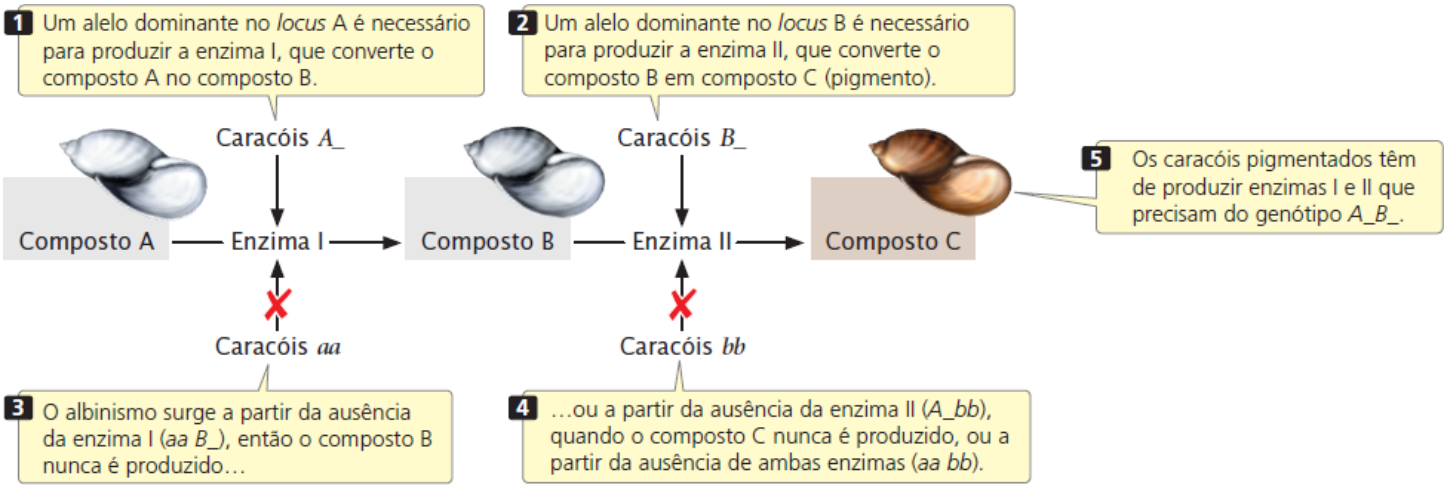


Figura 5.10 O pigmento é produzido em uma via de duas etapas nos caracóis.

A razão 9:7 observada nos caracóis de F_2 pode ser compreendida como uma modificação da razão 9:3:3:1, observada quando dois indivíduos heterozigotos para os dois *loci* são cruzados. A razão 9:7 surge quando os alelos dominantes em ambos os *loci* (*A_B_*) produzem caracóis pigmentados e qualquer outro genótipo produz caracóis albinos:

P: *aa BB* albino × *AA bb* albino

↓

F_1: *Aa Bb* pigmentados

↓ Intercruzamento

F_2: 9/16 *A_B_* pigmentados

3/16 *aa B_* albinos

3/16 *A_ bb* albinos } 7/16 albinos

1/16 *aa bb* albino

É provável que essa razão 9:7 seja produzida por uma via de duas etapas de produção de pigmentos (**Figura 5.10**). O pigmento (composto C) é produzido apenas após o composto A ter sido convertido no composto B pela enzima I e após o composto B ter sido convertido no composto C pela enzima II. Pelo menos um alelo A dominante no primeiro *locus* é necessário para produzir a enzima I; da mesma forma, pelo menos um alelo B dominante no segundo *locus* é necessário para produzir a enzima II. O albinismo surge a partir da ausência do composto C, que pode ocorrer de uma de três formas: primeiro, dois alelos recessivos no primeiro *locus* (genótipo *aa B_*) podem evitar a produção da enzima I, então o composto B nunca é produzido; segundo, os alelos recessivos no segundo *locus* (genótipo *A_ bb*) podem evitar a produção da enzima II, de modo que o composto B nunca é convertido no composto C; e, terceiro, dois alelos recessivos podem estar presentes em ambos os *loci* (*aa bb*), levando à ausência das enzimas I e II. Neste exemplo de interação gênica, *a* é epistático para *B* e *b* é epistático para *A*, e *ambos* são alelos epistáticos recessivos, pois a existência de duas cópias de um alelo *a* ou alelo *b* é necessária para suprimir a produção do pigmento. Este exemplo difere da supressão da cor do pelo no labrador retriever já que os alelos recessivos em cada um dos dois *loci* são capazes de suprimir a produção do pigmento nos caracóis, enquanto os alelos recessivos em um único *locus* suprimem a expressão do pigmento nos cães.

Conceitos

A epistasia é o disfarce da expressão de um gene por outro gene em um diferente *locus*. O gene epistático

faz o disfarce, o gene hipostático é mascarado. Os alelos epistáticos podem ser dominantes ou recessivos.

Checagem dos conceitos 7

Vários gatos brancos são cruzados e produzem os seguintes tipos de descendentes: 12/16 brancos, 3/16 pretos e 1/16 cinza. Qual é o genótipo dos descendentes pretos?

a. *Aa.*
b. *Aa Bb.*
c. *A_ B_.*
d. *A_ b.*

Conceitos conectantes

Como interpretar as razões produzidas pela interação gênica

Várias razões modificadas que surgem a partir da interação gênica estão apresentadas no **Quadro 5.2**. Cada um destes exemplos representa uma modificação da razão di-híbrida básica 9:3:3:1. Ao interpretar a base genética das razões modificadas, devemos ter vários pontos em mente. Primeiro, a herança dos genes que produzem essas características não é diferente da herança dos genes que codificam características genéticas simples. Os princípios de segregação e segregação independente de Mendel ainda se aplicam; cada organismo tem dois alelos em cada *locus*, que se separam na meiose, e os genes em diferentes *loci* se separam de forma independente. A única diferença é em como os *produtos* dos genótipos interagem para produzir o fenótipo. Assim, não podemos considerar a expressão dos genes em cada *locus* separadamente, mas sim considerar como os genes em diferentes *loci* interagem.

Um segundo ponto é que, nos exemplos que estudamos, as proporções fenotípicas eram sempre em 16 avos, em todos os cruzamentos, os pares de alelos se separam em dois *loci* com separação independente. A probabilidade de herdar um dos dois alelos em um *locus* é 1/2. Como existem dois *loci*, cada um com dois alelos, a probabilidade de herdar uma combinação específica de genes é $(1/2)^4 = 1/16$. Para os cruzamentos tri-híbridos, as proporções dos descendentes devem ser uma fração de 64 avos, porque $(1/2)^6 = 1/64$. Em geral, as proporções de descendentes devem ser frações de $(1/2)^{2n}$, em que n é igual ao número de *loci* com dois alelos que se separam no cruzamento.

Em raros casos, os cruzamentos não produzem exatamente 16 descendentes; portanto, as modificações de uma razão di-híbrida nem sempre são claras. É mais fácil observar as razões di-híbridas modificadas se o número de indivíduos de cada fenótipo for expresso na fração de 16 avos:

$$\frac{x}{16} = \frac{\text{Número de descendentes com um fenótipo}}{\text{Número total de descendentes}}$$

em que $x/16$ é igual à proporção de descendentes com um fenótipo específico. Se solucionarmos para x (a proporção do fenótipo específico na fração de 16), teremos:

$$x = \frac{\text{Número de descendentes com um fenótipo} \times 16}{\text{Número total de descendentes}}$$

Por exemplo, suponha que cruzamos dois homozigotos, intercruzamos os descendentes F_1 e obtemos 63 indivíduos vermelhos, 21 marrons e 28 brancos. Usando a fórmula anterior, descobrimos a razão fenotípica de F_1: vermelho = (63 × 16)/112 = 9; marrom = (21 × 16)/112 = 3; e branco = (28 × 16)/112 = 4. A razão fenotípica é 9:3:4.

Um ponto final a ser considerado é como designar os genótipos para os fenótipos nas razões modificadas que resultam da interação gênica. Não tente *memorizar* os genótipos associados a todas as razões modificadas apresentados no Quadro 5.2. Pelo contrário, pratique relacionar as razões modificadas com as conhecidas como a razão di-híbrida 9:3:3:1. Suponha que temos 15/16 de descendentes verdes e 1/16 de descendentes brancos em um cruzamento entre duas plantas. Se compararmos esta razão 15:1 com a razão di-híbrida padrão 9:3:3:1, veremos que 9/16 + 3/16 + 3/16 é igual a 15/16. Todos os genótipos associados a essas proporções no cruzamento di-híbrido (*A_ B_*, *A_ bb* e *aa B_*) têm de gerar o mesmo fenótipo, os descendentes verdes. O genótipo *aa bb* compõe 1/16 dos descendentes em um cruzamento di-híbrido, os

descendentes brancos neste cruzamento.

Ao designar os genótipos para os fenótipos nas razões modificadas, às vezes, os estudantes se confundem com quais letras usar para qual fenótipo. Suponha que obtenhamos a seguinte razão fenotípica: 9/16 pretos: 3/16 marrons: 4/16 brancos. Qual genótipo indicamos para os descendentes marrons, *A_ bb* ou *aa B_*? Ambas as respostas estão corretas porque as letras são apenas símbolos arbitrários para as informações genéticas. O que é importante perceber sobre esta razão é que o fenótipo marrom surge quando existem dois alelos recessivos em um *locus*.

Quadro 5.2 Razões fenotípicas di-híbridas modificadas por causa da interação gênica.

Razão*	Genótipo				Tipo de interação	Exemplo discutido no capítulo
	A_ B_	***A_ bb***	***aa B_***	***aa bb***		
9:3:3:1	9	3	3	1	Nenhuma	Formato e cor das ervilhas
9:3:4	9	3	4		Epistasia recessiva	Cor do pelo no labrador retriever
12:3:1	12		3	1	Epistasia dominante	Cor na abóbora
9:7	9	7			Epistasia recessiva duplicada	Albinismo nos caracóis
9:6:1	9	6		1	Interação duplicada	—
15:1	15			1	Epistasia dominante duplicada	—
13:3	13		3		Epistasia dominante e recessiva	—

*Cada razão é produzida por um cruzamento di-híbrido (*Aa Bb* × *Aa Bb*). As barras compartilhadas representam combinações de genótipos que geram o mesmo fenótipo.

Problema desenvolvido

Uma cepa homozigota de milho-amarelo é cruzada com uma cepa homozigota de milho-roxo. Os descendentes F_1 são intercruzados, produzindo 119 núcleos roxos e 89 núcleos amarelos (os descendentes). Qual é o genótipo dos núcleos amarelos?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

O genótipo dos núcleos amarelos.

Quais informações são fornecidas para solucionar o problema?

- Um milho homozigoto amarelo é cruzado com um milho homozigoto roxo
- O número de descendentes roxos e amarelos produzidos pelo cruzamento.

Etapas para a solução

Primeiro devemos considerar se o cruzamento entre as cepas amarela e roxa pode ser um cruzamento monoíbrido para um traço dominante simples, o que produziria uma razão 3:1 em F_2 (*Aa* × *Aa* → 3/4 *A_* e 1/4 *aa*).

De acordo com essa hipótese, esperaríamos 156 descendentes roxos e 52 descendentes amarelos:

Fenótipo	Genótipo	Número observado	Número esperado
Roxo	*A_*	119	3/4 × 208 = 156
Amarelo	*AA*	89	1/4 × 208 = 52
Total		208	

Observamos que os números esperados não se encaixam nos números observados. Se fizéssemos um teste qui-quadrado (ver *Capítulo 3*), obteríamos um valor de qui-quadrado calculado de 35,08, que tem a probabilidade muito menor que 0,05, indicando ser muito pouco provável que, quando esperamos uma razão 3:1, obtenhamos 119 descendentes roxos e 89 descendentes amarelos. Portanto, podemos rejeitar a hipótese de que estes resultados foram produzidos por um cruzamento monoíbrido.

Outra hipótese possível é que os descendentes F_2 observados são uma razão 1:1. Entretanto, aprendemos no Capítulo 3 que uma razão 1:1 é produzida por um cruzamento entre um heterozigoto e um homozigoto (*Aa* × *aa*) e, a partir das informações fornecidas, o cruzamento não foi entre um heterozigoto e um homozigoto, porque as cepas originais eram homozigotas. Além disso, um teste qui-quadrado comparando os números observados com uma razão esperada 1:1 gera um valor de qui-quadrado calculado de 4,32, que tem uma probabilidade de menos de 0,05.

A seguir, devemos ver se os resultados podem ser explicados por um cruzamento di-híbrido (*Aa Bb* × *Aa Bb*). Um cruzamento di-híbrido resulta em proporções fenotípicas na fração de 16 avos. Podemos aplicar a fórmula indicada anteriormente neste capítulo para determinar o número da fração de 16 avos para cada fenótipo:

$$x = \frac{\text{número de descendentes com um fenótipo} \times 16}{\text{número total de descendentes}}$$

$$x_{(\text{roxo})} = \frac{119 \times 16}{208} = 9{,}15$$

$$x_{(\text{amarelo})} = \frac{89 \times 16}{208} = 6{,}85$$

Assim, os descendentes roxos e amarelos surgem em uma razão aproximada de 9:7.

Podemos testar essa hipótese com um teste qui-quadrado:

Fenótipo	Genótipo	Número observado	Número esperado
Roxo	?	119	9/16 × 208 = 117
Amarelo	?	89	7/16 × 208 = 91
Total		208	

$$\chi^2 = \sum \frac{(\text{observado} - \text{esperado})^2}{\text{esperado}}$$

$$= \frac{(119 - 117)^2}{117} + \frac{(89 - 91)^2}{91}$$

$$= 0{,}034 + 0{,}044 = 0{,}078$$

$$\text{Grau de liberdade} = n - 1 = 2 - 1 = 1$$

$$P > 0{,}05$$

A probabilidade associada ao valor de qui-quadrado é maior que 0,05, indicando que existe um bom ajuste entre os resultados observados e uma razão 9:7.

Agora precisamos determinar como um cruzamento di-híbrido pode produzir uma razão 9:7 e quais genótipos correspondem aos dois fenótipos. Um cruzamento di-híbrido sem epistasia produz uma razão 9:3:3:1:

$$Aa\ Bb \times Aa\ Bb$$

$$\downarrow$$

$A\_\ B\_$	9/16
$A\_\ bb$	3/16
$aa\ B\_$	3/16
$aa\ bb$	1/16

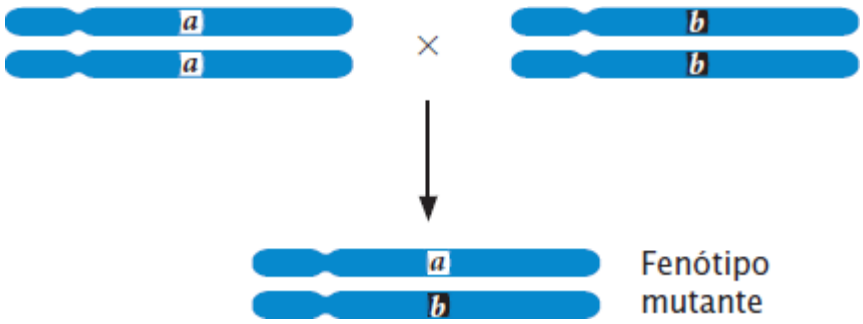
Como 9/16 dos descendentes do cruzamento do milho são roxos, esta cor tem de ser produzida pelos genótipos *A_ B_*; em outras palavras, os núcleos que têm pelo menos um alelo dominante no primeiro *locus* e um no segundo são roxos. As proporções de todos os outros genótipos (*A_ bb*, *aa B_* e *aa bb*) é 7/16, que é a proporção de descendentes no cruzamento do milho que são amarelos, então qualquer núcleo que não tenha um alelo dominante no primeiro e no segundo *locus* é amarelo.

Agora teste sua compreensão sobre epistasia ao fazer o **Problema 26** no final do capítulo.

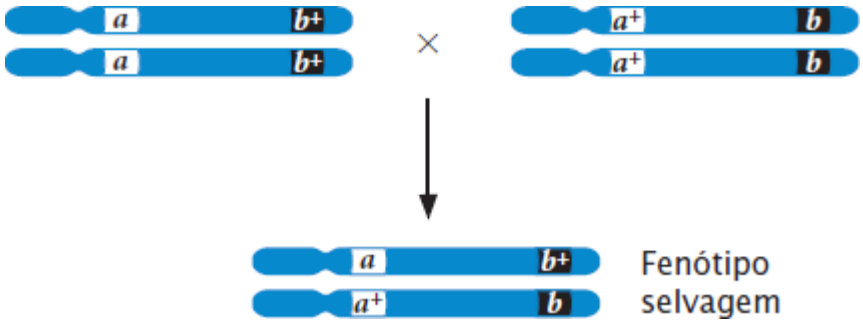
Complementação | Como determinar se as mutações estão no mesmo *locus* ou em *loci* diferentes

Como sabemos se diferentes mutações que afetam uma característica ocorrem no mesmo *locus* (são alélicas) ou em diferentes *loci*? Nas moscas-da-fruta, por exemplo, o *branco* é uma mutação recessiva ligada ao X que produz olhos brancos em vez dos olhos vermelhos encontrados na variante selvagem; *damasco* é uma mutação recessiva ligada ao X que produz olhos de cor laranja-claro. As mutações de *branco* e *damasco* ocorrem no mesmo *locus* ou em *loci* diferentes? Podemos usar o teste de complementação para responder a esta pergunta.

Para fazer um **teste de complementação** em mutações recessivas, os genitores homozigotos para diferentes mutações são cruzados, produzindo descendentes que são heterozigotos. Se as mutações são alélicas (ocorrem no mesmo *locus*), então os descendentes heterozigotos têm apenas os alelos mutantes (*a b*) e exibem um fenótipo mutante:



Se, por outro lado, as mutações ocorrem em diferentes *loci*, cada um dos genitores homozigotos tem genes selvagens em outro *locus* (*aa* b^+b^+ e a^+a^+ *bb*); então os descendentes heterozigotos herdam um alelo mutante e um alelo selvagem em cada *locus*. Nesse caso, a existência de um alelo selvagem complementa a mutação de cada *locus* e os descendentes heterozigotos têm o fenótipo selvagem:



A **complementação** ocorre se um organismo com duas mutações recessivas tem um fenótipo selvagem, indicando que as

mutações são genes não alélicos. Ocorre uma falta de complementação quando duas mutações recessivas ocorrem no mesmo *locus*, produzindo um fenótipo mutante.

Quando o teste de complementação é aplicado nas mutações *branco* e *damasco*, todos os descendentes heterozigotos têm olhos claros, demonstrando que os olhos brancos e os olhos damasco são produzidos por mutações que ocorrem no mesmo *locus* e são alélicos.

Conceitos

Um teste de complementação é usado para determinar se duas mutações ocorrem no mesmo *locus* (são alélicas) ou em diferentes *loci*.

Checagem dos conceitos 8

O rajado (aparência de tigre) é um traço recessivo nos buldogues e chihuahuas. Quais tipos de cruzamentos você faria para determinar se os genes para *rajado* nos buldogues e chihuahuas estão no mesmo *locus*?

A complexa genética da cor do pelo nos cães

A genética da cor do pelo nos cães é um excelente exemplo de como as interações complexas entre os genes participam na determinação de um fenótipo. Os cães domésticos têm diferentes formatos, tamanhos e cores. Há milhares de anos, as pessoas vêm cruzando os cães por causa de traços específicos, produzindo um grande número de tipos que observamos hoje em dia. Cada raça de cão carreia uma seleção de alelos a partir do *pool* de genes do cão ancestral que definem as características de uma raça. O genoma do cão doméstico foi completamente sequenciado em 2004, o que facilitou muito o estudo da genética canina.

Vamos considerar quatro *loci* na lista a seguir que são importantes para produzir muitas das diferenças que observamos na cor e padrão entre as raças de cães. Ao interpretar a base genética das diferenças na cor do pelo dos cães, consideramos como a expressão de um gene particular é modificada pelos efeitos de outros genes. Tenha em mente que os *loci* adicionais não listados aqui podem modificar as cores produzidas por estes quatro *loci* e que nem todos os geneticistas concordam com a genética de variação de cor em algumas raças.

1. *Locus* Agouti (***A***). Este *locus* tem cinco alelos comuns que determinam a profundidade e a distribuição de cor no pelo de um cão:
   - A^s pigmento preto uniforme.
   - a^w Agouti, ou cinza de lobo. Os pelos codificados por este alelo têm uma aparência “salpicada”, produzida por uma banda de pigmento amarelo no pelo preto.
   - a^y Amarelo. O pigmento preto está muito reduzido, então o pelo todo é amarelo.
   - a^s marcas de sela (preto nas costas, com extensas marcas de bege na cabeça e pernas).
   - a^t bicolor (cor preta sobre quase todo corpo com marcas de bege nos pés e sobrancelhas).

Os alelos A^s e a^y são, em geral, dominantes sobre os outros alelos, mas as relações de dominância são complexas e não estão totalmente compreendidas.

2. *Locus* **(*B*) preto**. Este *locus* determina se o pigmento preto pode ser formado. A cor verdadeira do pelo de um cão depende dos efeitos dos genes em outros *loci* (como nos *loci A* e *E*). Dois alelos são comuns:
   - *B* permite que o pigmento preto seja produzido.
   - *b* O pigmento preto não pode ser produzido e os cães pigmentados podem ser chocolate, castanhos, bege ou vermelhos.

   O alelo *B* é dominante sobre o alelo *b*.
3. *Locus* **(*E*) de extensão**. Quatro alelos neste *locus* determinam onde o genótipo no *locus A* é expresso. Por exemplo, se um cão tem o alelo A^s (preto uniforme) no *locus A*, então o pigmento preto será estendido por todo o pelo ou pode ser restrito a algumas áreas, dependendo dos alelos presentes no *locus E*. As áreas onde o *locus A* não é expresso podem surgir como amarelas, vermelhas ou bege, dependendo se há genes particulares em outros *loci*. Quando A^s está no *locus A*, os quatro alelos no *locus E* têm os seguintes efeitos:
   - E^m Máscara preta com pelo bege.
   - *E* O *locus A* é expresso por todo o corpo (preto uniforme)
   - e^{br} Rajado, no qual preto e amarelo estão em camadas, semelhante ao observado em tigres.
   - e Sem preto no pelo, mas o nariz e olhos podem ser pretos.

   As relações de dominância entre estes alelos são pouco conhecidas.

4. *Locus* **(*S*) manchado**. Os alelos neste *locus* determinam se haverá manchas brancas. Existem quatro alelos comuns:

S	Sem manchas.
s^i	Mancha irlandesa, numerosas manchas brancas.
s^p	Mancha piebald, diferentes manchas de branco.
s^w	Piebald branco extremo, quase todo branco.

O alelo *S* é completamente dominante sobre os alelos s^i, s^p e s^w; os alelos s^i e s^p são dominantes sobre o alelo s^w ($S > s^i$, $s^p > s^w$). A relação entre s^i e s^p não está bem definida, na verdade, eles podem não ser alelos separados. Os genes em outros *loci* pouco conhecidos também modificam o padrão de manchas.

Para ilustrar como os genes nestes *loci* interagem para determinar a cor do pelo de um cão, vamos considerar alguns exemplos.

Labrador retriever. Este cão (**Figura 5.11 A**) pode ser preto, marrom ou amarelo. A maior parte deles é homozigota A^sA^s *SS*, assim, eles variam apenas nos *loci B* e *E*. O alelo A^s permite que o pigmento preto seja expresso, se um cão é preto depende de quais genes são encontrados nos *loci B* e *E*. Como já foi discutido no início deste capítulo, todos os labradores pretos devem ter pelo menos um alelo *B* e um alelo E (*B*_ *E*_). Os cães marrons são *bb* homozigotos e têm pelo menos um alelo *E* (*bb E*_). Os cães amarelos são um resultado da existência de *ee* (*B*_ *ee* ou bb ee). Os labradores retrievers são homozigotos para o alelo *S*, que produz uma cor uniforme, as poucas manchas brancas que aparecem em alguns cães desta raça são consequência de outros genes modificadores.

Beagle. A maior parte dos beagles é a^sa^s *BB* s^Ps^P homozigota, embora outros alelos nesses *loci* estejam presentes em alguns casos. O alelo a^s produz as marcas em sela – costas e lateral escuras com cabeça e pernas claras, que são características da raça. O alelo *B* permite que a cor preta seja produzida, mas sua distribuição é limitada pelo alelo a^s. Muito dos beagles é *E*_, mas o genótipo *ee* surge em alguns casos, gerando alguns beagles beges. As manchas brancas nos beagles são causadas pelo alelo s^P.

Dálmata. Os dálmatas (**Figura 5.11 B**) têm uma interessante composição genética. A maior parte é homozigota A^sA^s *EE* s^ws^w, então eles variam apenas no *locus B*. Observe que estes cães têm o genótipo A^sA^s *EE*, que permite um pelo uniforme que seria preto, se o genótipo *B*_ estiver presente, ou marrom (chamado de esverdeado), se o genótipo *bb* estiver presente. Entretanto, a existência do alelo s^w produz um pelo branco, mascarando a expressão da cor uniforme. A cor do cão aparece apenas nas manchas pigmentadas, geradas pela existência de um alelo em outro *locus* que possibilita que a cor penetre em um número limitado de manchas.

O **Quadro 5.3** fornece os genótipos mais comuns de outras raças de cães. **Resolva o Problema 33**

5.3 O sexo influencia a herança e a expressão dos genes de várias maneiras

No Capítulo 4, estudamos as características codificadas por genes localizados nos cromossomos sexuais (traços ligados ao sexo) e como sua herança difere da herança dos traços codificados por genes autossômicos. Os traços ligados ao X, por exemplo, são transmitidos do pai para a filha, mas nunca do pai para o filho, e os traços ligados ao Y são transmitidos de pai para todos os filhos do sexo masculino. Agora, vamos examinar influências adicionais do sexo, incluindo o efeito do sexo de um organismo na expressão de genes nos cromossomos autossômicos, nas características determinadas pelos genes localizados no citoplasma e nas características para as quais o genótipo apenas da genitora determina o fenótipo dos descendentes. Finalmente, observamos situações nas quais a expressão dos genes nos cromossomos autossômicos é afetada pelo sexo do genitor do qual os genes são herdados.

Figura 5.11 A cor do pelo nos cães é determinada pelas interações gênicas em vários *loci*. A. A maior parte dos labradores retrievers tem o genótipo A^sA^s *SS*, variando apenas nos *loci B* e *E*. **B.** Os dálmatas têm genótipo A^sA^s *EE* s^ws^w, variando no *locus B*, que torna os cães pretos (*B_*) ou marrons (*bb*). (*Parte A*: imagebroker/Alamy. *Parte B*: PhotoDisc.)

Quadro 5.3 Genótipos comuns em diferentes raças de cães.

Raça	Genótipos homozigotos comuns*	Outros alelos presentes dentro da raça
Basset hound	*BB EE*	a^y, a^tS, s^P, s^i
Beagle	a^sa^s *BB* s^ps^p	*E, e*
Buldogue inglês	*BB*	A^s, a^y, a^tE^m, *E*, $e^{br}S$, s^i, s^p, s^w
Chihuahua	*BB EE*	A^s, a^y, a^s, a^tB, *b* E^m, *E*, e^{br}, *e S*, s^i, s^p, s^w
Collie		a^y, a^ts^i, s^w
Dálmata	A^sA^s *EE* s^ws^w	*B, b*
Dobermann	a^ta^t *EE SS*	*B, b*
Pastor-alemão	*BB SS*	a^y, *a*, a^s, a^tE^m, *E, e*
Golden retriever	A^sA^s *BB SS*	*E, e*
Galgo	*BB*	A^s, a^yE, e^{br}, *e S*, s^p, s^w, s^i
Setter irlandês	*BB ee SS*	*A*, a^t
Labrador retriever	A^sA^s *SS*	*B, b E, e*
Poodle	*SS*	A^s, a^tB, *b E, e*
Rottweiler	a^ta^t *BB EE SS*	
São-bernardo	a^ya^y *BB*	E^m, *E* s^i, s^p, s^w

*A maior parte dos cães na raça é homozigota para esses genes, alguns poucos cães têm outros alelos nesses *loci*. Fonte: Dados de M. B. Willis, Genetics of the Dog (London: Witherby, 1989).

Características influenciadas e características limitadas pelo sexo

As **características influenciadas pelo sexo** são determinadas por genes autossômicos e herdadas de acordo com os princípios de Mendel, mas elas são expressas de forma diferente em machos e fêmeas. Neste caso, um determinado traço é expresso sem dificuldade em um sexo; em outras palavras, o traço tem maior penetrância em um dos sexos.

Por exemplo, a existência de barba em alguns bodes é determinada por um gene autossômico (B^b) que é dominante nos

machos e recessivo nas fêmeas. Nos machos, um único alelo é necessário para a expressão do traço, de modo que o homozigoto (B^bB^b) e o heterozigoto (B^bB^+) têm barba, enquanto o macho B^+B^+ não tem.

Genótipo	Machos	Fêmeas
B^+B^+	Sem barba	Sem barba
B^+B^b	Com barba	Sem barba
B^bB^b	Com barba	Com barba

Ao contrário, a fêmea precisa de dois alelos de barba para que este traço seja expresso: o homozigoto B^bB^b tem barba, enquanto o heterozigoto (B^bB^+) e o homozigoto (B^+B^+) não têm.

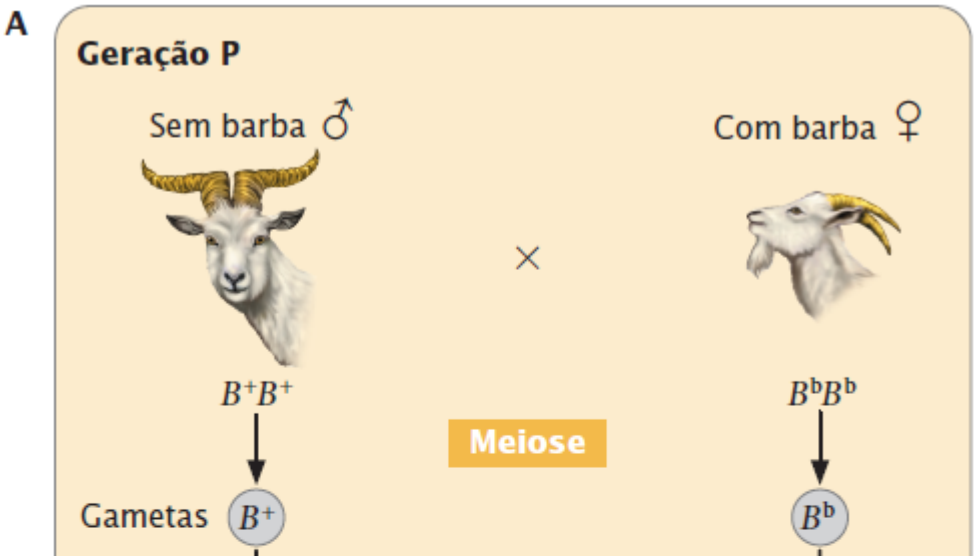
A chave para compreender a expressão do gene da barba é examinar o heterozigoto. Nos machos (para os quais a existência de barba é dominante), o genótipo heterozigoto produz barba, mas nas fêmeas (para as quais a ausência de barba é dominante), o genótipo heterozigoto produz uma cabra sem barba.

A **Figura 5.12 A** ilustra um cruzamento entre um macho sem barba (B^+B^+) e uma fêmea com barba (B^bB^b). Os alelos se separam em gametas de acordo com o princípio de segregação de Mendel e todos os F_1 são heterozigotos (B^+B^b). Como o traço é dominante nos machos e recessivo nas fêmeas, todos os machos de F_1 terão barba e todas as fêmeas de F_1 não. Quando os descendentes F_1 são cruzados entre si, 1/4 dos descendentes é B^bB^b, 1/2 é B^bB^+ e 1/4 é B^+B^+ (**Figura 5.12 B**). Como os machos heterozigotos têm barba, 3/4 dos machos em F_2 têm barba, como as fêmeas heterozigotas não têm barba, apenas 1/4 das fêmeas de F_2 tem barba.

Uma forma extrema de herança influenciada pelo sexo, a **característica limitada pelo sexo** é codificada por genes autossômicos que são expressos em apenas um sexo; assim, o traço tem penetrância zero no outro sexo. Nas galinhas domésticas, alguns machos apresentam um padrão de plumagem chamado de penas em franjas (**Figura 5.13 A).** Outros machos e todas as fêmeas apresentam um padrão chamado de penas de galinha (**Figuras 5.13 B**). As penas em franjas são um traço autossômico recessivo limitado ao sexo para os machos. Como o traço é autossômico, os genótipos dos machos e fêmeas são os mesmos, mas os fenótipos produzidos por estes genótipos são diferentes nos machos e fêmeas:

Genótipo	Fenótipo do macho	Fenótipo da fêmea
HH	Penas de galinha	Penas de galinha
Hh	Penas de galinha	Penas de galinha
hh	Penas em franjas	Penas de galinha

Um exemplo de característica limitada pelo sexo nos seres humanos é a puberdade precoce limitada nos homens. Existem vários tipos de puberdade precoce nos seres humanos, a maior parte não genética. Essa puberdade nos homens, entretanto, é o resultado de um alelo autossômico dominante (*P*) que é expresso apenas nos homens, sendo as mulheres com esse gene de fenótipo normal. Os homens com puberdade precoce entram na puberdade cedo, em geral, antes dos 4 anos. Nesse momento, o pênis aumenta de tamanho, a voz engrossa e os pelos pubianos surgem. Não há comprometimento da função sexual; os homens acometidos são férteis, mas a maior parte é de adultos de baixa estatura porque os ossos longos param de crescer após a puberdade.



Fertilização

Geração F_1

Com barba ♂ Sem barba ♀

B^+B^b B^+B^b

B

Geração F_1

Com barba ♂ × Sem barba ♀

B^+B^b B^+B^b

Meiose

Gametas B^+ B^b B^+ B^b

Fertilização

Geração F_2

Sem barba ♂	Com barba ♂	Com barba ♂	Sem barba ♀	Sem barba ♀	Com barba ♀
1/4 B^+B^+	1/2 B^+B^b	1/4 B^bB^b	1/4 B^+B^+	1/2 B^+B^b	1/4 B^bB^b
	3/4				1/4

Conclusão: 3/4 dos machos têm barba, 1/4 das fêmeas tem barba.

Figura 5.12 Os genes que codificam os traços influenciados pelo sexo são herdados de acordo com os princípios de Mendel, mas expressos de forma diferente nos machos e fêmeas.

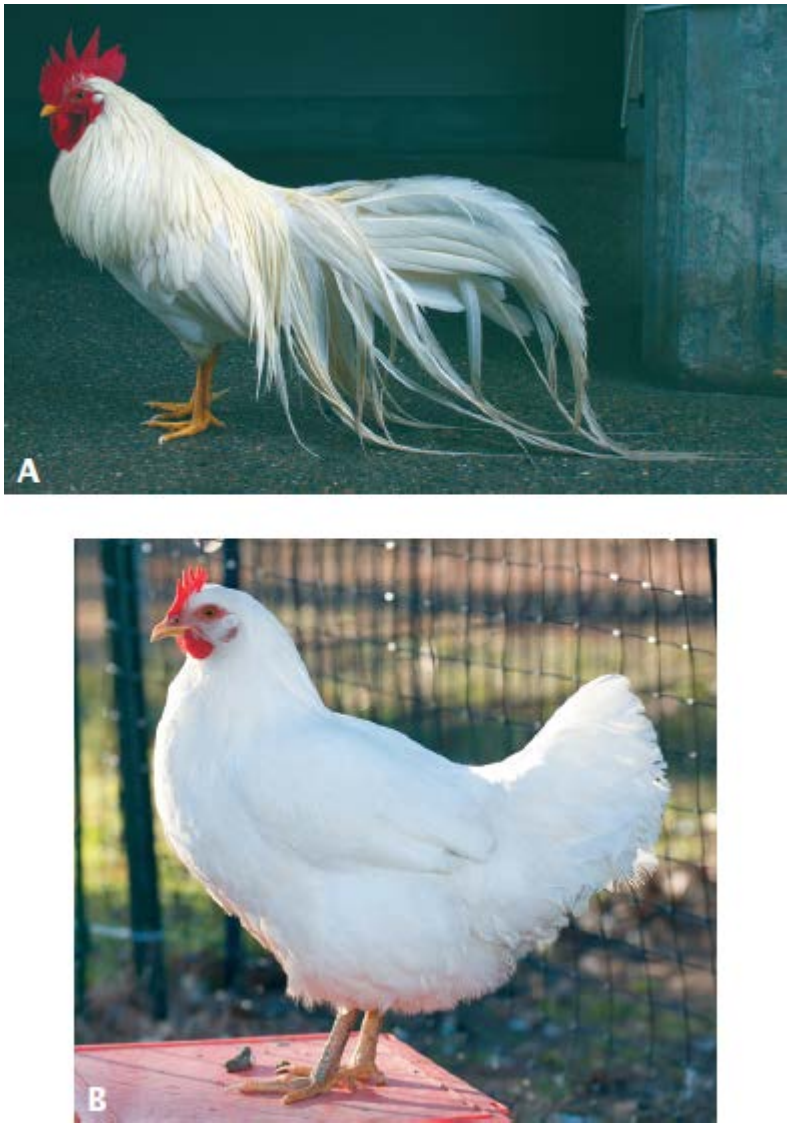


Figura 5.13 Uma característica limitada ao sexo é codificada por genes autossômicos que são expressos em apenas um sexo. Um exemplo são as penas em franjas nas galinhas, um traço autossômico recessivo limitado aos machos. **A.** Macho com penas em franja. **B.** Fêmea com penas de galinha. (*Parte A*: superstarjet/Getty Images. *Parte B*: Guy Sagi/istockphoto.)

Como esse traço é raro, em geral, os homens acometidos são heterozigotos (*Pp*). Um homem com puberdade precoce que engravida uma mulher sem história familiar da doença transmitirá o alelo para a puberdade precoce a 1/2 de seus filhos (**Figura 5.14 A**), mas apenas seus filhos do sexo masculino expressarão a característica. Se uma das filhas heterozigotas (*Pp*) engravidar de um homem com puberdade normal (*pp*), 1/2 de seus filhos exibirá a puberdade precoce (**Figura 5.14 B**). Dessa forma, uma característica limitada ao sexo pode ser herdada do pai ou da mãe, embora o traço apareça em apenas um sexo. **Resolva o Problema 35**

Conceitos

As características influenciadas pelo sexo são codificadas por genes autossômicos que são expressos mais facilmente em um sexo. As características limitadas ao sexo são codificadas por genes autossômicos cuja expressão está limitada a um sexo.

Checagem dos conceitos 9

Qual é a diferença das características influenciadas pelo sexo e limitadas pelo sexo das características ligadas ao sexo?

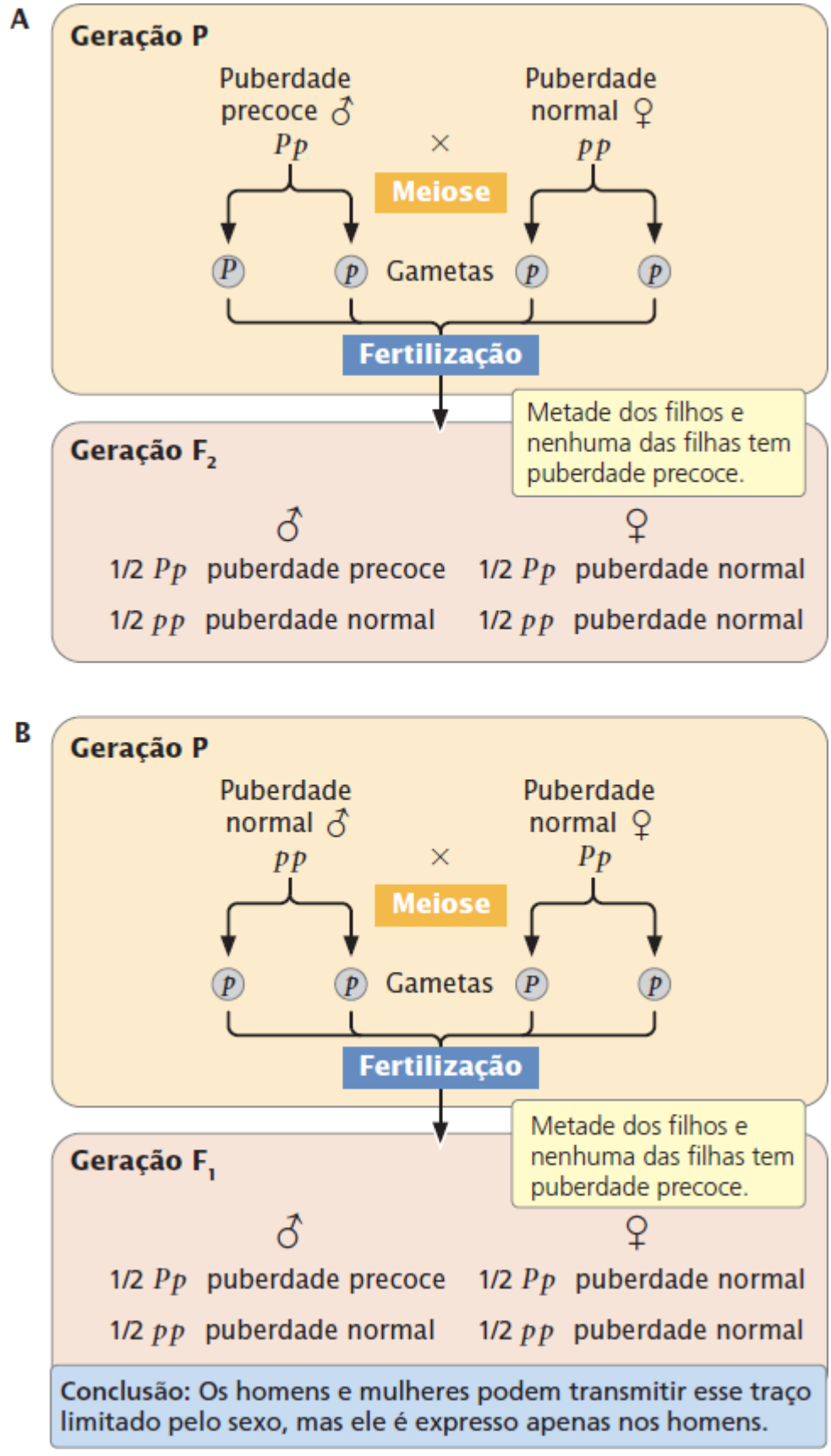


Figura 5.14 As características limitadas pelo sexo são herdadas de acordo com os princípios de Mendel. A puberdade precoce é um traço autossômico dominante limitado pelo sexo para os machos.

Herança citoplasmática

Os princípios de segregação e segregação independente de Mendel se baseiam na suposição de que os genes estão localizados em cromossomos dentro do núcleo da célula. Essa suposição é válida para grande parte das características genéticas e os princípios de Mendel permitem que possamos prever os tipos de descendentes que serão produzidos em um cruzamento genético. Entretanto, nem todo material genético de uma célula é encontrado no núcleo, com algumas características codificadas por genes localizados no citoplasma, características que exibem a **herança citoplasmática**.

Algumas organelas, em especial os cloroplastos e as mitocôndrias, têm DNA. O genoma mitocondrial humano tem cerca de 15.000 nucleotídios de DNA, que codificam 37 genes. Comparados com o DNA nuclear, com três bilhões de nucleotídios, que codificam 20.000 genes, o tamanho do genoma mitocondrial é muito pequeno, contudo, os genes mitocondriais e cloroplásticos codificam algumas características importantes. Os detalhes moleculares do DNA extranuclear são discutidos no Capítulo 11; vamos nos concentrar nos *padrões* da herança citoplasmática.

A herança citoplasmática é diferente da herança de características codificadas pelos genes nucleares em várias questões importantes. Um zigoto herda os genes nucleares de ambos os genitores, mas, normalmente, todas as suas organelas citoplasmáticas e seus genes citoplasmáticos vêm de apenas um dos gametas; em geral, o óvulo. Um espermatozoide do genitor contribui apenas com um conjunto de genes nucleares. Assim a maior parte dos traços de herança citoplasmática ocorre em machos e fêmeas e é transmitida da mãe para os descendentes, nunca do pai. Os cruzamentos recíprocos, portanto, geram resultados diferentes quando os genes citoplasmáticos codificam um traço. Em alguns organismos, entretanto, os genes citoplasmáticos são herdados do genitor masculino apenas ou de ambos.

As características de herança citoplasmática frequentemente exibem grande variação fenotípica porque nenhum mecanismo análogo à mitose ou meiose garante que estes genes sejam igualmente distribuídos na divisão celular. Assim, diferentes células e descendentes terão várias proporções de genes citoplasmáticos.

Considere os genes mitocondriais. A maior parte das células tem milhares de mitocôndrias e cada uma tem de 2 a 10 cópias do DNA mitocondrial (mtDNA). Suponha que metade das mitocôndrias em uma célula tenha uma cópia selvagem normal de mtDNA e a outra metade tenha uma cópia mutante (**Figura 5.15**). Na divisão celular, a separação das mitocôndrias para as células-filhas é aleatória. Por acaso, uma célula pode receber a grande parte do mtDNA mutante e outra célula pode receber principalmente o mtDNA selvagem. Dessa maneira, diferentes descendentes da mesma mãe e até células em um mesmo descendente podem ter diferentes fenótipos. Os traços codificados pelo DNA do cloroplasto (cpDNA) são variáveis do mesmo modo. As características que os traços de herança citoplasmática exibem estão resumidas no **Quadro 5.4**.

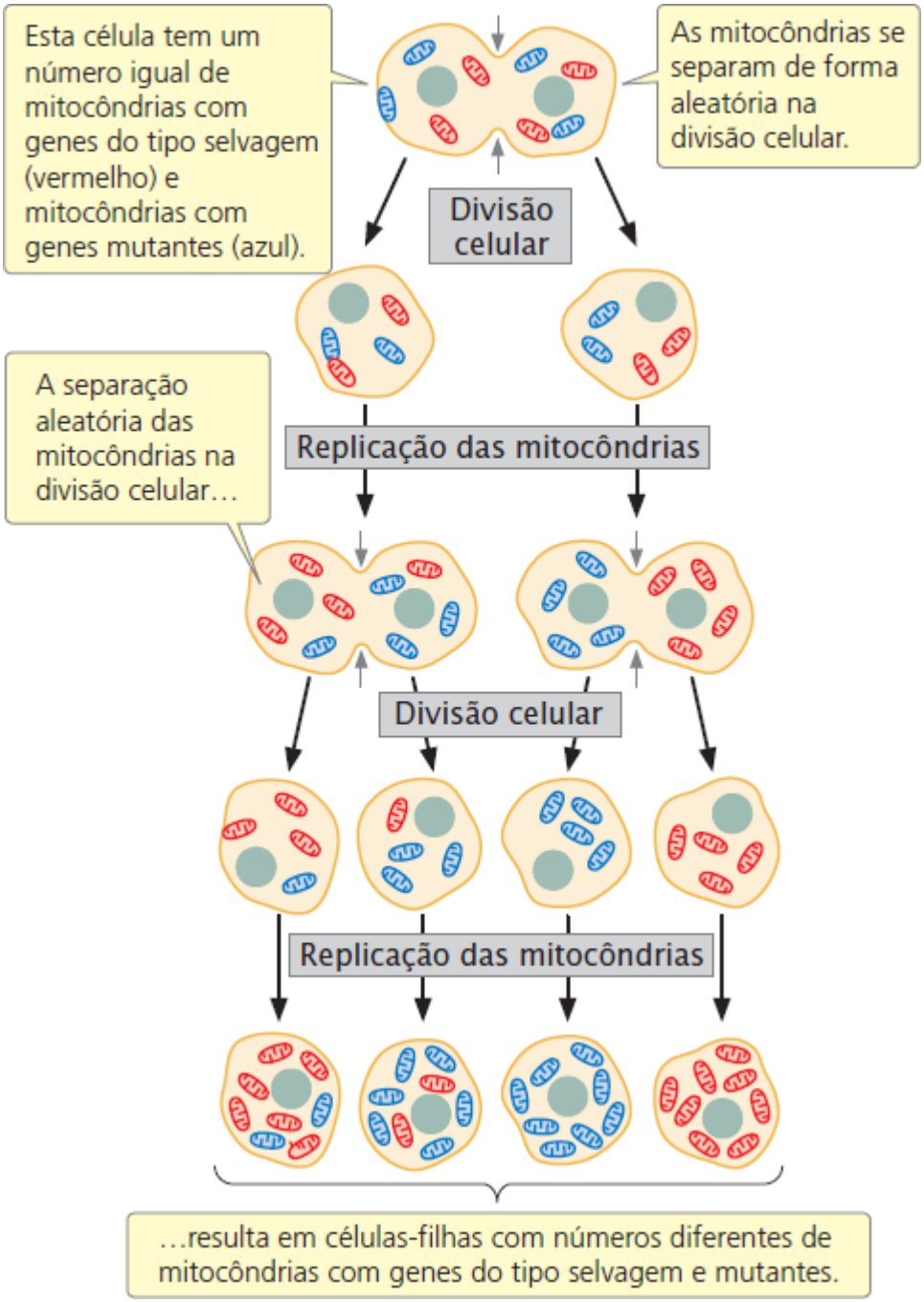


Figura 5.15 As características de herança citoplasmática exibem grande variação fenotípica porque as células e os indivíduos têm várias proporções de genes citoplasmáticos.

Quadro 5.4 Características dos traços herdados do citoplasma.

1. Presente em machos e fêmeas.
2. Geralmente herdado de um genitor, em geral a genitora.
3. Cruzamentos recíprocos produzem resultados diferentes.
4. Exibem grande variação fenotípica, mesmo dentro de uma família.

Variegação na planta maravilha. Em 1909, a herança citoplasmática foi reconhecida por Carl Correns como uma exceção aos princípios de Mendel. Correns, um dos biólogos que redescobriu o trabalho de Mendel, estudou a herança da variegação da folha da planta maravilha, *Mirabilis jalapa*. Correns descobriu que as folhas e brotos de uma variedade da maravilha eram variegadas, apresentando uma mistura de manchas verdes e brancas. Ele também observou que alguns ramos da cepa variegada tinham folhas verdes sem manchas e outros tinham folhas brancas sem manchas. Cada ramo produziu flores, então Correns pôde cruzar as flores dos ramos variegados, verdes e brancos em todas as combinações (**Figura 5.16**). As sementes a partir dos ramos verdes sempre geravam descendentes verdes, não importa se o pólen era de um ramo verde, branco ou variegado. Da

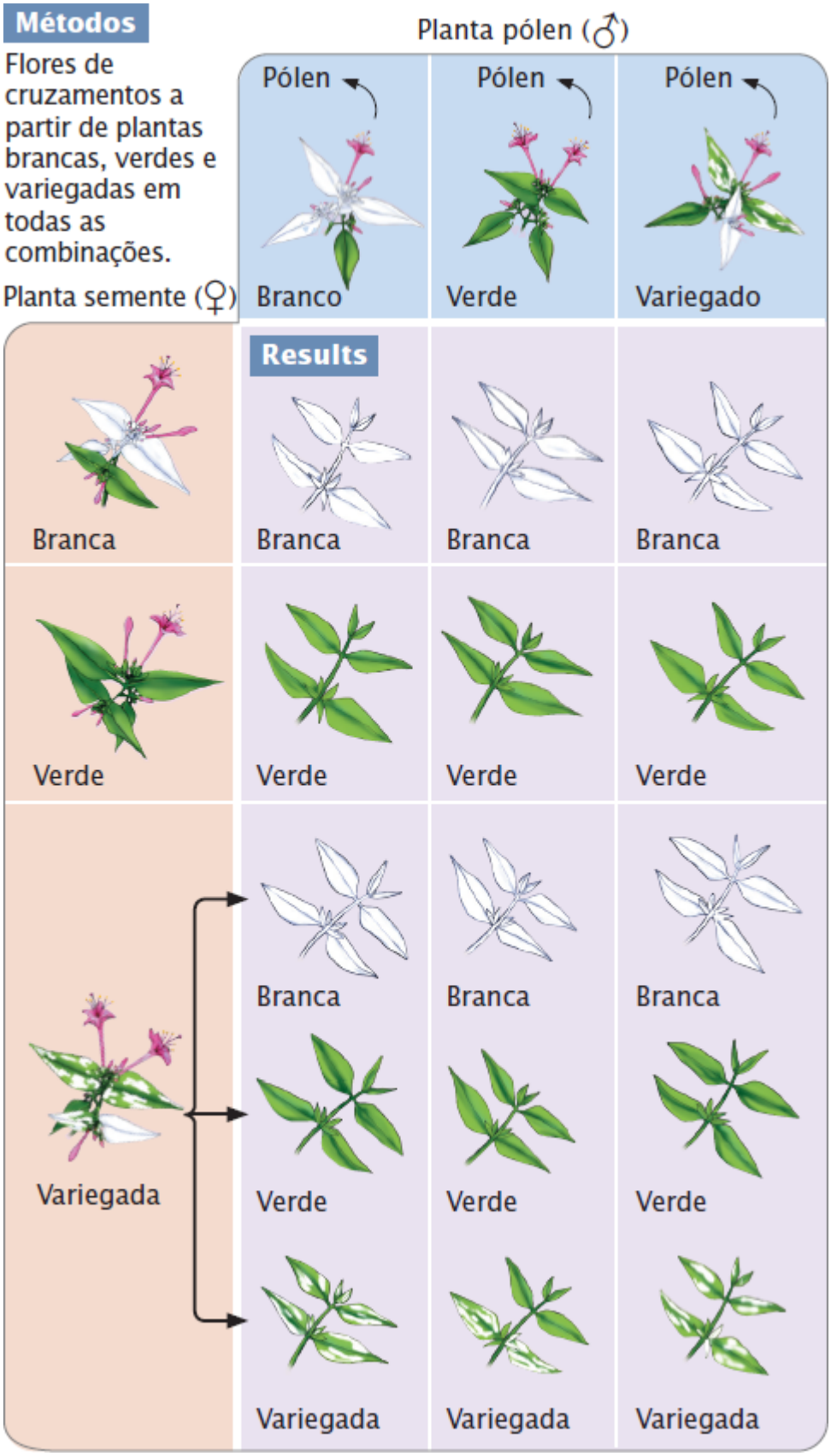
mesma forma, as flores nos ramos brancos sempre produziam descendentes com flores brancas. As flores nos ramos variegados geravam descendentes verdes, brancos e variegados, sem uma razão específica.

Os cruzamentos de Correns demonstraram a herança citoplasmática da variegação da planta maravilha. Os fenótipos dos descendentes eram totalmente determinados pela genitora, nunca pelo genitor (a fonte do pólen). Além disso, a produção de todos os três fenótipos pelas flores nos ramos variegados é compatível com a herança citoplasmática. A variegação nessas plantas é causada por um gene defeituoso no cpDNA, que leva à não produção do pigmento verde clorofila. As células dos ramos verdes têm apenas cloroplastos normais, as dos ramos brancos têm apenas cloroplastos anormais e as dos variegados têm uma mistura dos dois tipos de cloroplastos. Nas plantas dos ramos variados, a separação aleatória dos cloroplastos durante a ovogênese produz alguns óvulos com cpDNA normal, que se desenvolvem nos descendentes verdes, outros óvulos com cpDNA anormal que apenas se desenvolvem em descendentes brancos, e finalmente, ainda outros óvulos com uma mistura de cpDNA normal e anormal que se desenvolvem nos descendentes variegados.

Doenças mitocondriais. Já foram identificadas várias doenças nos seres humanos (muitas raras) que exibem herança citoplasmática. Esses distúrbios surgem a partir de mutações no mtDNA, nos genes que codificam componentes da cadeia de transporte de elétrons, que gera a maior parte do ATP (trifosfato de adenosina) na respiração celular aeróbica. Uma delas é a neuropatia óptica hereditária de Leber (LHON). Pacientes com esta doença apresentam perda rápida e bilateral da visão, resultado da morte das células no nervo óptico. Essa perda de visão ocorre no início da vida adulta (geralmente entre 20 e 24 anos), mas pode surgir em qualquer momento após a adolescência. Existe grande variabilidade clínica na gravidade da doença, mesmo em determinada família. A neuropatia óptica hereditária de Leber exibe herança citoplasmática: o traço é transmitido da mãe para todas as crianças, filhos e filhas, igualmente.

Experimento

Pergunta: Como é a herança da cor do caule e da folha na planta maravilha?



Conclusão: O fenótipo da prole é determinado pelo fenótipo do ramo a partir do qual a semente se originou, não a partir do ramo no qual o pólen se originou. A cor do caule e da folha exibe herança citoplasmática.

Figura 5.16 Os cruzamentos para o tipo de folha na planta maravilha ilustram a herança citoplasmática.

Efeito genético materno

Um fenômeno genético que às vezes é confundido com a herança citoplasmática é o **efeito genético materno**, no qual o fenótipo dos descendentes é determinado pelo genótipo da mãe. Na herança citoplasmática, os genes para uma característica são herdados apenas de um genitor, em geral da mãe. No efeito genético materno, os genes são herdados de ambos os genitores,

mas o fenótipo dos descendentes não é determinado pelo seu próprio genótipo, mas pelo de sua mãe.

Esse efeito surge, com frequência, quando substâncias presentes no citoplasma de um óvulo (codificadas pelos genes nucleares da genitora) são essenciais no início do desenvolvimento. Um excelente exemplo é a enroscadura do casco do caracol *Lymnaea peregra* (**Figura 5.17**), descrito na introdução deste capítulo. A enroscadura do casco para a direita é chamada destra, enquanto a enroscadura para esquerda é chamada canhota. No *Lymnaea peregra*, o sentido da enroscadura é determinado por um par de alelos, o alelo para destro (s^+) é dominante sobre o alelo para canhoto (s). Entretanto, o sentido da enroscadura é determinado não pelo genótipo do próprio caracol, mas pelo genótipo de sua *genitora*. O sentido da enroscadura é influenciado pela forma na qual o citoplasma se divide logo após a fertilização, que por vez, é determinado por uma substância produzida pela genitora e transmitida para os descendentes no citoplasma do óvulo.

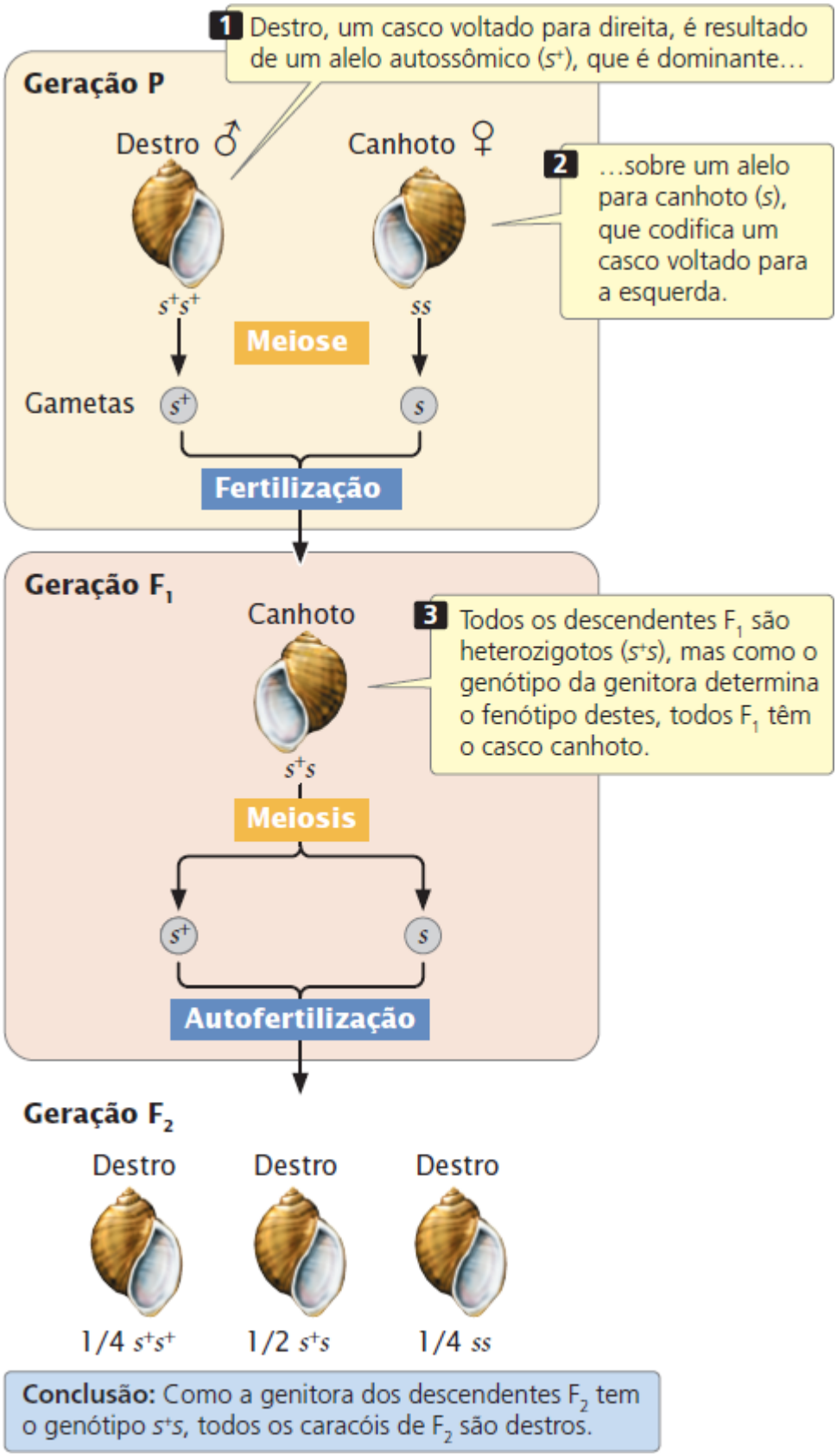


Figura 5.17 No efeito genético materno, o genótipo da genitora determina o fenótipo dos descendentes. A enroscadura do casco de um caracol é um traço que exibe o efeito genético materno.

Se um macho homozigoto para os alelos destro (s^+s^+) for cruzado com uma fêmea homozigota para os alelos canhoto (*ss*), todos os descendentes F_1 serão heterozigotos (s^+s) e terão o casco com enroscadura para a esquerda, porque o genótipo da genitora (*ss*) codifica esse tipo de enroscadura (ver Figura 5.17). Se os caracóis F_1 se autofertilizarem, a razão genotípica de F_2 é 1 s^+s^+: 2 s^+s: 1 *ss*.

Observe que o fenótipo de todos os caracóis de F_2 é destro, independentemente de seus genótipos. Os descendentes F_2 são destros porque o genótipo de sua genitora (s^+s) codifica um casco com enroscadura para a direita e determina o fenótipo. Com o efeito genético materno, o fenótipo dos descendentes não é necessariamente o mesmo do fenótipo da genitora, porque o fenótipo dos descendentes é determinado pelo *genótipo* da genitora, não pelo fenótipo. Nenhum genótipo do genitor ou dos próprios descendentes influencia o fenótipo dos descendentes. Entretanto, um macho influencia o fenótipo da geração F_2: ao contribuir com os genótipos de suas descendentes fêmeas, ele afeta os fenótipos de seus descendentes. Os genes que exibem o efeito genético materno são, portanto, transmitidos dos machos para as futuras gerações. Ao contrário, os genes que exibem herança citoplasmática são transmitidos apenas por um dos sexos (em geral, as fêmeas). **Resolva o Problema 39**

Conceitos

As características que exibem herança citoplasmática são codificadas por genes presentes no citoplasma e em geral são herdadas de um genitor, mais comumente da genitora. No efeito genético materno, o genótipo da genitora determina o fenótipo dos descendentes.

Checagem dos conceitos 10

Como você pode determinar se um traço específico é gerado pela herança citoplasmática ou pelo efeito genético materno?

Imprinting genômico

Um princípio básico da genética mendeliana é que a origem do genitor de um gene não afeta sua expressão e, portanto, os cruzamentos recíprocos geram resultados idênticos. Observamos que existem algumas características genéticas – as codificadas por genes ligados ao X e genes citoplasmáticos – para as quais os cruzamentos recíprocos não geram os mesmos resultados. Nesses casos, os machos e as fêmeas não contribuem com o mesmo material genético para os descendentes. Com relação aos genes autossômicos, os machos e fêmeas contribuem com o mesmo número de genes e se considera que os genes paternos e maternos tenham efeitos iguais. Entretanto, a expressão de alguns genes é afetada de maneira significativa pela origem dos genitores. Esse fenômeno, a expressão diferencial do material genético dependendo se ele foi herdado do genitor ou genitora, é chamado de ***imprinting* genômico**.

Um gene que exibe *imprinting* genômico nos camundongos e seres humanos é o *Igf2*, que codifica uma proteína chamada de fator de crescimento semelhante à insulina tipo 2 (Igf2). Os descendentes herdam um alelo de *Igf2* de sua genitora e um de seu genitor. A cópia paterna do *Igf2* é expressa de forma ativa no feto e na placenta, mas a cópia materna é completamente silenciada (**Figura 5.18**). Os descendentes machos e fêmeas têm os genes de *Igf2*, e a chave para saber se o gene é expresso é o sexo do genitor que transmitiu o gene. No exemplo apresentado, o gene é expresso apenas quando é transmitido pelo genitor masculino. Em outros traços de *imprinting* genômico, o traço é expresso apenas quando o gene é transmitido pela genitora. De uma forma que não é compreendida, o alelo *Igf2* paterno (mas, não a cópia materna) promove o crescimento placentário e fetal; quando a cópia paterna de *Igf2* é eliminada nos camundongos, surgem uma pequena placenta e descendentes recém-nascidos com baixo peso.

O *imprinting* genômico está envolvido em várias doenças nos seres humanos, incluindo as síndromes de Prader-Willi e Angelman. As crianças com a síndrome de Prader-Willi têm pés e mãos pequenos, baixa estatura, pouco desenvolvimento sexual e incapacidade intelectual. Essas crianças são pequenas quando nascem e se amamentam mal, mas, quando crianças, desenvolvem apetite voraz e frequentemente se tornam obesos. Muitas crianças com a síndrome de Prader-Willi não têm uma pequena região do braço longo do cromossomo 15. A deleção dessa região é sempre herdada do *pai*. Assim, as crianças com a síndrome de Prader-Willi não têm a cópia paterna dos genes no braço longo do cromossomo 15.

A deleção dessa mesma região do cromossomo 15 também pode ser herdada da *mãe*, mas essa herança causa um conjunto completamente diferente de sintomas, produzindo a síndrome de Angelman. As crianças com a síndrome de Angelman exibem riso frequente, movimento muscular não controlado, uma grande boca e convulsões incomuns. As crianças não têm uma cópia materna dos genes no braço longo do cromossomo 15. Para que ocorra o desenvolvimento normal, são necessárias as cópias

dessa região do cromossomo 15 de ambos os pais.

Muitos genes que sofrem *imprinting* nos mamíferos estão associados ao crescimento fetal. *Imprinting* também foi observado nas plantas, com a expressão diferencial dos genes paternos e maternos no endosperma, que, como a placenta nos mamíferos, fornece nutrientes para o crescimento do embrião. Os mecanismos de *imprinting* ainda estão em investigação, mas a metilação do DNA – o acréscimo de grupos metila (CH_3) aos nucleotídios do DNA (ver *Capítulos 10* e *17*) – é essencial para o processo. Nos mamíferos, a metilação é apagada nas células germinativas de cada geração e, então, restabelecida no curso de formação do gameta, com os espermatozoides e óvulos sofrendo diferentes níveis de metilação, que levam à expressão diferenciada dos alelos do genitor e da genitora nos descendentes. Algumas das diferentes formas nas quais o sexo interage com a hereditariedade estão resumidos no **Quadro 5.5**.

Quadro 5.5 As influências do sexo na hereditariedade.

Fenômeno genético	Fenótipo determinado por
Característica ligada ao sexo	Genes localizados no cromossomo sexual
Característica influenciada pelo sexo	Genes em cromossomos autossômicos que são mais facilmente expressos em um sexo
Característica limitada pelo sexo	Genes autossômicos cuja expressão é limitada por um sexo
Efeito genético materno	Genótipo nuclear da genitora
Herança citoplasmática	Genes citoplasmáticos que, em geral, são completamente herdados de apenas um genitor
Imprinting genômico	Genes cuja expressão é afetada pelo sexo do genitor que o transmite

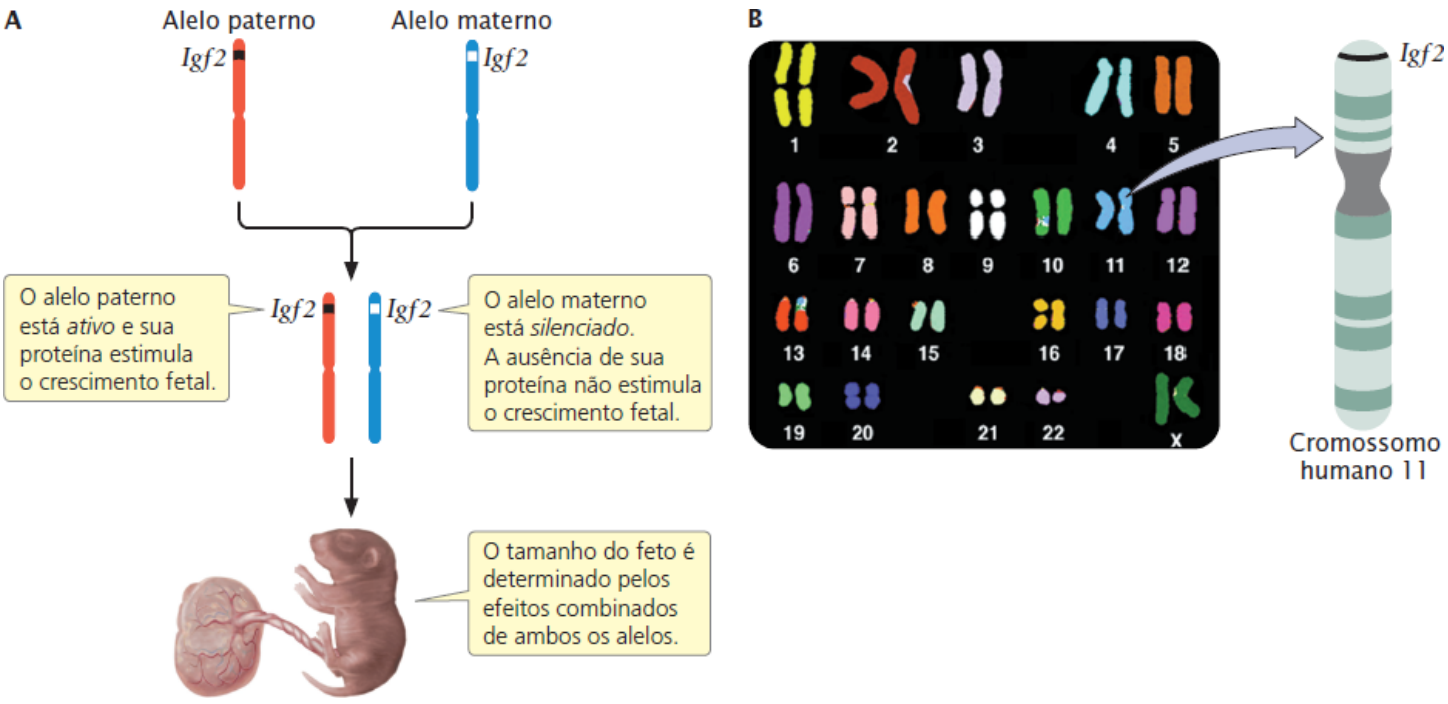


Figura 5.18 O *imprinting* genômico do gene *Igf2* nos camundongos e nos seres humanos afeta o crescimento fetal. A. O alelo *Igf2* paterno está ativo no feto e na placenta, enquanto o alelo materno está silencioso. **B.** O *locus Igf2* humano está no braço curto do cromossomo 11, o *locus* nos camundongos está no cromossomo 7. (Cortesia de Dr. Thomas Ried e Dr. Evelin Schrock.)

Epigenética. O *imprinting* genômico é apenas uma forma de um fenômeno conhecido como **epigenética**. A maior parte dos traços é codificada por informações genéticas contidas na sequência das bases de nucleotídios do DNA, o código genético, que será discutido no Capítulo 15. Entretanto, alguns traços são causados por alteração no DNA, como o acréscimo de grupos metila a algumas bases de DNA (metilação de DNA) que afetam a forma na qual as sequências de DNA são expressas. Essas

mudanças são estáveis e podem ser herdadas porque são transmitidas de uma célula para outra.

No *imprinting* genômico, o fato de o gene ser transmitido pelo óvulo ou pelo espermatozoide determina quanto de metilação do DNA ocorre. O padrão de metilação em um gene é copiado quando o DNA é replicado e, portanto, permanece no gene à medida que ele é transmitido entre as células por meio da mitose. Entretanto, o padrão de metilação pode ser modificado ou removido quando o DNA é transmitido para um gameta. Um gene metilado no espermatozoide pode ser desmetilado quando é transmitido para um óvulo da filha. No final, a quantidade de metilação determina se o gene é expresso nos descendentes.

Esses tipos de mudanças reversíveis ao DNA que influenciam a expressão dos traços são chamados de marcadores epigenéticos. A inativação de um dos cromossomos X nas fêmeas dos mamíferos (discutido no *Capítulo 4*) é outro tipo de mudança epigenética. Vamos considerar as mudanças epigenéticas com mais detalhes no Capítulo 21.

Conceitos

No *imprinting* genômico, a expressão de um gene é influenciada pelo sexo do genitor que transmite o gene para os descendentes. Os marcadores epigenéticos são mudanças reversíveis no DNA que não alteram a sequência de bases, mas podem afetar como um gene é expresso.

Checagem dos conceitos 11

Que tipo de marcador genético é responsável pelo *imprinting* genômico?

5.4 A antecipação é a expressão mais forte ou precoce dos traços nas gerações seguintes

Outro fenômeno genético que não é explicado pelos princípios de Mendel é a **antecipação**, na qual um traço genético é expresso mais intensamente ou é expresso de forma precoce à medida que é passado de geração para geração. No início de 1900, vários médicos observaram que muitos pacientes com distrofia miotônica moderada a grave – uma doença muscular autossômica dominante – tinham ancestrais com a forma leve da doença. Essas observações levaram ao conceito de antecipação. Entretanto, o conceito rapidamente foi desacreditado pelos geneticistas porque não havia mecanismo claro para explicá-lo. A genética tradicional afirmava que os genes eram passados inalterados dos pais para os descendentes e os geneticistas tendiam a atribuir a antecipação a viés observacional.

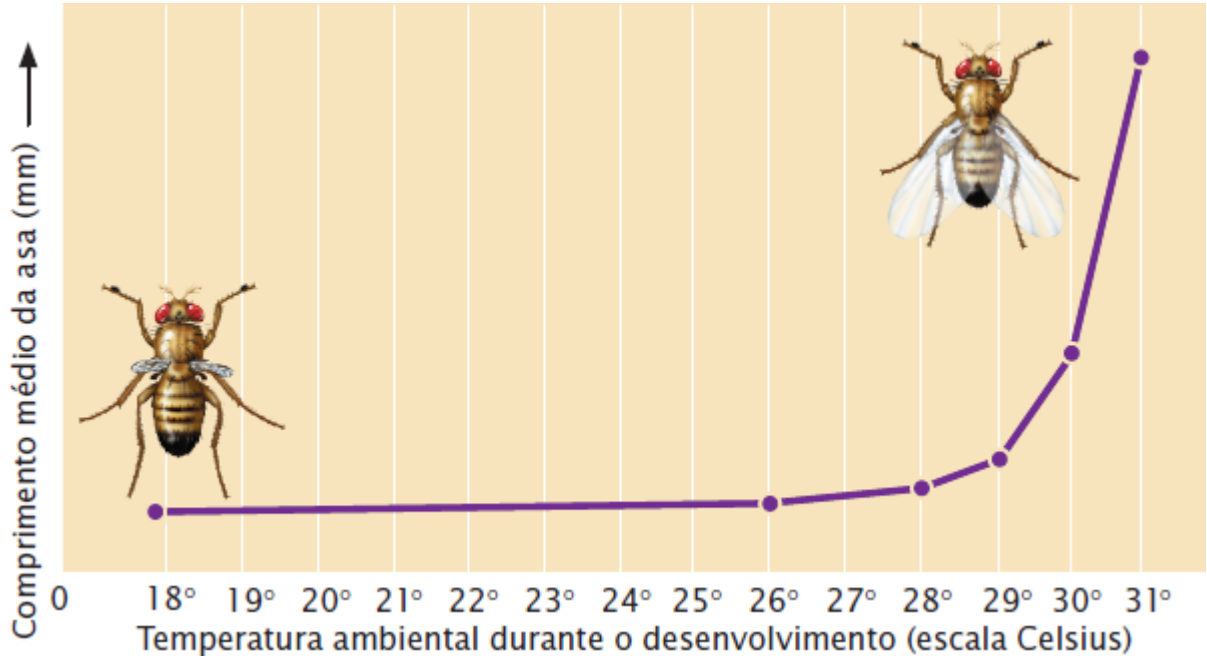


Figura 5.19 A expressão da mutação vestigial na *Drosophila* é dependente da temperatura. Quando crescem em temperaturas abaixo de 29°C, as moscas homozigotas para asa vestigial têm suas asas muito reduzidas; já em temperaturas acima de 31°C, elas desenvolvem asas normais. (Dados de M. H. Harnly, Journal of Experimental Zoology 56:363-379, 1936.)

Atualmente, a pesquisa restabeleceu a antecipação como um legítimo fenômeno genético. A mutação que causa a distrofia miotônica é uma região instável de DNA que pode aumentar em tamanho à medida que o gene é transmitido de uma geração

para outra. A idade do início e a gravidade da doença estão correlacionadas com o tamanho da região instável, de modo que um aumento no tamanho da região através das gerações produz a antecipação. O fenômeno agora está envolvido em várias doenças genéticas. Vamos examinar esses tipos interessantes de mutações com mais detalhes no Capítulo 18.

Conceitos

A antecipação é a expressão mais forte ou precoce de um traço genético nas gerações seguintes. É causada por uma região instável do DNA que aumenta de tamanho de uma geração para outra.

5.5 A expressão de um genótipo pode ser afetada por efeitos ambientais

No Capítulo 3 aprendemos que cada fenótipo é o resultado de um genótipo que se desenvolve dentro de um ambiente específico, e cada genótipo pode produzir vários fenótipos diferentes, dependendo das condições ambientais nas quais o desenvolvimento ocorre. Por exemplo, uma mosca-da-fruta homozigota para a mutação vestigial (*vg vg*) desenvolve asas reduzidas quando alcança temperaturas abaixo de 29°C, mas o mesmo genótipo desenvolve asas muito maiores quando alcança 31°C (**Figura 5.19**).

Para a maior parte das características discutidas até o momento, o efeito do ambiente no fenótipo é desprezível. As ervilhas de Mendel com genótipo *yy*, por exemplo, desenvolviam sementes verdes independentemente do ambiente no qual elas eram cultivadas. Da mesma forma, pessoas com o genótipo I^AI^A têm o antígeno A nas suas hemácias independentemente da sua dieta, situação socioeconômica ou ambiente familiar. Para outros fenótipos, entretanto, os efeitos ambientais têm um papel mais importante.

Efeitos ambientais no fenótipo

A expressão fenotípica de alguns genótipos depende, de modo crítico, de um ambiente específico. Por exemplo, o alelo *himalaio* nos coelhos produz o pelo escuro nas extremidades do corpo – nariz, orelhas e patas. O pigmento escuro se desenvolve, entretanto, apenas quando um coelho é criado em uma temperatura de 25°C ou menos; se um coelho himalaio for criado a 30°C, não surgem manchas escuras. A expressão do alelo *himalaio* é dependente da temperatura; uma enzima necessária para a produção do pigmento escuro é inativada a altas temperaturas. O pigmento é restrito ao nariz, às patas e às orelhas de um coelho himalaio porque a temperatura corporal do cerne do animal normalmente está acima de 25°C e a enzima só funciona nas células de extremidades relativamente frias. O alelo *himalaio* é um exemplo de **alelo sensível à temperatura**, um alelo cujo produto é funcional apenas em algumas temperaturas. Da mesma forma, as asas vestigiais na *Drosophila melanogaster* são causadas por uma mutação dependente da temperatura (ver Figura 5.19).

Os fatores ambientais também são importantes na expressão de várias doenças genéticas nos seres humanos. A fenilcetonúria (PKU) ocorre por causa de um alelo autossômico recessivo que provoca comprometimento intelectual. A doença surge a partir de um defeito em uma enzima que normalmente metaboliza o aminoácido fenilalanina. Quando essa enzima é defeituosa, a fenilalanina não é metabolizada e seu acúmulo provoca dano cerebral nas crianças. Uma simples mudança do ambiente, ou seja, a instituição de uma dieta de baixo teor de fenilalanina para o recém-nascido, evita a ocorrência de comprometimento intelectual.

Esses exemplos ilustram o ponto no qual os genes e seus produtos não agem de maneira isolada; ao contrário, eles frequentemente interagem com fatores ambientais. Em alguns casos, os fatores ambientais sozinhos podem produzir um fenótipo igual ao produzido por um genótipo, chamado de **fenocópia**. Nas moscas-da-fruta, por exemplo, a mutação autossômica recessiva *sem olhos* produz olhos muito reduzidos. O fenótipo sem olhos também pode ser produzido ao expor as larvas das moscas normais a metaborato de sódio.

Conceitos

A expressão de muitos genes é modificada pelo ambiente. Uma fenocópia é um traço produzido por efeitos ambientais que mimetiza o fenótipo produzido por um genótipo.

Checagem dos conceitos 12

Como você pode determinar se um fenótipo como os olhos reduzidos nas moscas-da-fruta é causado por uma mutação recessiva ou é uma fenocópia?

A herança das características contínuas

Até o momento, tratamos principalmente das características que têm apenas alguns fenótipos diferentes. Nas ervilhas de Mendel, por exemplo, as sementes eram lisas ou rugosas, amarelas ou verdes. Já os pelos dos cães eram pretos, marrons ou amarelos, e os tipos sanguíneos eram A, B, AB ou O. Tais características, que têm fenótipos de fácil distinção, são chamadas de **características descontínuas**.

Nem todas as características exibem fenótipos descontínuos. A altura do ser humano é um exemplo dessa característica; as pessoas não têm alturas diferentes, mas apresentam uma ampla faixa de altura. De fato, existem tantos possíveis fenótipos da altura do humano que é preciso usar uma medida para descrever a altura de uma pessoa. As características que exibem uma distribuição contínua de fenótipos são chamadas de **características contínuas**. Como estas características têm muitos possíveis fenótipos e têm de ser descritas em termos quantitativos, as características contínuas também são chamadas de **características quantitativas**.

As características contínuas frequentemente surgem porque genes em muitos *loci* interagem para produzir os fenótipos. Quando um único *locus* com dois alelos codifica uma característica, existem três genótipos possíveis: *AA*, *Aa* e *aa*. Com dois *loci*, cada um com dois alelos, existem $3^2 = 9$ possíveis genótipos. O número de genótipos que codificam uma característica é 3^n, em que *n* é igual ao número de *loci*, cada um com dois alelos, que influenciam a característica. Por exemplo, quando uma característica é determinada por oito *loci*, cada um com dois alelos, existem $3^8 = 6.561$ possíveis genótipos diferentes para ela. Se cada genótipo produz um fenótipo diferente, vários fenótipos serão possíveis. As sutis diferenças entre os fenótipos serão indistinguíveis, e a característica parece ser contínua. As características codificadas por genes em muitos *loci* são chamadas de **características poligênicas**.

O oposto da poligenia é a **pleiotropia**, na qual um gene afeta múltiplas características. Muitos genes exibem pleiotropia. A fenilcetonúria, já mencionada, é resultado de um alelo recessivo. As pessoas homozigotas para este alelo, se não forem tratadas, apresentam déficit intelectual, olhos azuis e pele de cor clara. O alelo letal que gera a cor amarela do pelo nos camundongos também é pleiotrópico. Além da sua letalidade e do efeito na cor do pelo, o gene provoca uma doença semelhante ao diabetes melito e maior propensão para desenvolver tumores.

Frequentemente, os fenótipos de características contínuas também são influenciados por fatores ambientais. Cada genótipo é capaz de produzir uma gama de fenótipos. Nesta situação, o fenótipo específico resultante depende do genótipo e das condições ambientais nas quais ele se desenvolve. Por exemplo, apenas três genótipos codificam uma característica, mas, como cada genótipo produz uma gama de fenótipos associados a diferentes ambientes, o fenótipo da característica exibe uma distribuição contínua. Muitas características contínuas são poligênicas e influenciadas por fatores ambientais, sendo chamadas de **características multifatoriais**, porque muitos fatores ajudam a determinar o fenótipo.

A herança das características contínuas parece complexa, mas os alelos em cada *locus* seguem os princípios de Mendel e são herdados da mesma forma que os alelos que codificam características simples, descontínuas. Entretanto, como participam muitos genes, os fatores ambientais influenciam o fenótipo, e como os fenótipos não se segregam em tipos diferentes, não podemos observar as razões distintas que permitem que interpretemos a base genética das características descontínuas. Para analisar as características contínuas, é preciso usar ferramentas estatísticas especiais, como será discutido no Capítulo 24.

Resolva o Problema 45

Conceitos

Características descontínuas exibem alguns fenótipos diferentes; as características contínuas exibem uma gama de fenótipos. Uma característica contínua é frequentemente produzida quando os genes em muitos *loci* e os fatores ambientais se combinam para determinar um fenótipo.

Checagem dos conceitos 13

Qual é a diferença entre poligenia e pleiotropia?

Resumo dos conceitos

- A dominância sempre se refere aos genes no mesmo *locus* (genes alélicos) e pode ser compreendida sobre como o fenótipo do heterozigoto se relaciona com os fenótipos dos homozigotos
- A dominância é completa quando um heterozigoto tem o mesmo fenótipo de um homozigoto, é incompleta quando o heterozigoto tem um fenótipo intermediário entre os fenótipos dos dois genitores homozigotos, e é codominante quando o heterozigoto exibe traços de ambos os genitores homozigotos

- O tipo de dominância não afeta a herança de um alelo, mas a expressão fenotípica dele. A classificação da dominância depende do nível do fenótipo examinado
- A penetrância é a porcentagem de indivíduos com um genótipo específico que expressa o fenótipo associado. A expressividade é o grau no qual um traço é expresso
- Os alelos letais provocam a morte de um organismo, em geral no estágio inicial do desenvolvimento, e modificam as razões fenotípicas
- Os alelos múltiplos se referem à existência de mais de dois alelos em um *locus* de um grupo. Sua presença aumenta o número de genótipos e fenótipos possíveis
- A interação gênica se refere à interação entre genes em diferentes *loci* para produzir um único fenótipo. Um gene epistático em um *locus* suprime ou mascara a expressão dos genes hipostáticos em outros *loci*. A interação gênica produz razões fenotípicas que são modificações de razões di-híbridas
- As características influenciadas pelo sexo são codificadas por genes autossômicos que são expressos mais facilmente em um sexo. As características limitadas pelo sexo são codificadas por genes autossômicos que são expressos em apenas um sexo
- Na herança citoplasmática, os genes para a característica são encontrados em organelas e, em geral, são herdados de um único genitor (tipicamente, materno). O efeito genético materno ocorre quando um descendente herda genes de ambos os genitores, mas os genes nucleares da genitora determinam o seu fenótipo
- O *imprinting* genômico se refere a características codificadas por genes autossômicos cuja expressão é afetada pelo sexo do genitor que transmite os genes. Os efeitos epigenéticos como o *imprinting* genômico são causados por alteração ao DNA – como a metilação do DNA – que não afetam a sequência de bases do DNA
- A antecipação se refere a um traço genético que é expresso mais intensamente ou surge de forma precoce nas gerações seguintes
- Os fenótipos são frequentemente modificados por fatores ambientais. Uma fenocópia é um traço produzido por efeitos ambientais que mimetiza o fenótipo produzido por um genótipo
- As características contínuas são as que exibem uma ampla faixa de fenótipos, e são frequentemente produzidas pelos efeitos combinados de muitos genes e de efeitos ambientais.

Termos importantes

Alelo letal
Alelo sensível à temperatura
Alelos múltiplos
Antecipação
Característica contínua
Característica descontínua
Característica influenciada pelo sexo
Característica limitada pelo sexo
Característica multifatorial
Característica poligênica
Característica quantitativa
Codominância
Complementação
Dominância completa
Dominância incompleta
Efeito genético materno
Epigenética
Epistasia
Expressividade
Fenocópia
Gene epistático
Gene hipostático
Herança citoplasmática
Heterozigoto composto
Imprinting genômico
Interação gênica
Penetrância

Penetrância incompleta
Pleiotropia
Teste de complementação

Respostas da Checagem dos conceitos

1. b.
2. Na dominância completa, o heterozigoto expressa o mesmo fenótipo que um dos homozigotos. Na dominância incompleta, o heterozigoto tem um fenótipo que é intermediário entre dois homozigotos. Na codominância, o heterozigoto tem um fenótipo que expressa simultaneamente os fenótipos de ambos homozigotos.
3. d.
4. b.
5. c.
6. A interação gênica é a interação entre genes em diferentes *loci*. A dominância é a interação entre genes em um único *locus*.
7. d.
8. Cruzar um buldogue homozigoto para *rajado* com um chihuahua homozigoto para *rajado*. Se os dois genes para *rajado* são alélicos, todos os descendentes serão rajados: $bb \times bb \rightarrow$ todos bb (rajados). Se, por outro lado, o rajado nas duas raças é causado por genes recessivos em diferentes *loci*, nenhum dos descendentes será rajado: $a^+a^+\ bb \times aa\ b^+b^+ \rightarrow a^+a\ b^+b$.
9. Tanto os traços influenciados por sexo como os limitados por sexo são codificados por genes autossômicos cuja expressão é afetada pelo sexo do genitor que tem os genes. As características ligadas ao sexo são determinadas pelos genes nos cromossomos sexuais.
10. Os traços de herança citoplasmática são codificados por genes existentes no citoplasma, que, em geral, são herdados apenas da genitora. Portanto, um traço de herança citoplasmática será sempre transmitido pelas fêmeas. Os traços causados pelo efeito genético materno são codificados por genes autossômicos e podem, portanto, ser transmitidos pelos machos, embora nenhum traço do organismo seja determinado pelo genótipo da genitora.
11. Metilação do DNA.
12. O cruzamento de duas moscas sem olhos e o cruzamento de uma mosca sem olhos com uma mosca selvagem. Crie os descendentes de ambos os cruzamentos no mesmo ambiente. Se o traço for consequente a uma mutação recessiva, todos os descendentes das duas moscas sem olhos não terão olhos, enquanto pelo menos algumas das descendentes das moscas sem olhos e selvagens devem ser selvagens. Se o traço for consequente a uma fenocópia, não devem existir diferenças nos descendentes dos dois cruzamentos.
13. A poligenia se refere à influência dos múltiplos genes na expressão de uma única característica. A pleiotropia se refere ao efeito de um único gene na expressão de múltiplas características.

Problemas desenvolvidos

Problema 1

Um geneticista cruzou dois camundongos amarelos com pelo liso e obteve os seguintes descendentes.

1/2 amarelo, liso

1/6 amarelo, encrespado

1/4 cinza, liso

1/2 cinza, encrespado

a. Forneça uma explicação genética para os resultados e indique os genótipos dos genitores e descendentes deste cruzamento.
b. Quais cruzamentos adicionais podem ser feitos para confirmar se sua explicação está correta?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

a. Uma explicação genética para a herança da cor e tipo de pelo nos camundongos. Os genótipos dos genitores.

b. Exemplos de outros cruzamentos que podem ser feitos para determinar se a explicação fornecida em (a) está correta.

Quais informações são fornecidas para solucionar o problema?

- Os fenótipos dos genitores
- Fenótipos e proporções de diferentes tipos de descendentes.

Para a solução deste problema, revise:
Alelos letais, na Seção 5.1, e razões para cruzamentos genéticos simples (Quadro 3.5).

Etapas para a solução

a. Este cruzamento se preocupa com duas características separadas – a cor e o tipo de pelo. Primeiro, vamos examinar a herança da cor. Dois camundongos amarelos são cruzados, produzindo 1/2 + 1/6 = 3/6 + 1/6 = 2/3 de camundongos amarelos e 1/4 + 1/12 = 3/12 + 1/12 = 4/12 = 1/3 de camundongos cinza. Aprendemos no Capítulo 3 que uma razão 2:1 é produzida quando existe um gene letal recessivo:

Dica: Examine as razões de descendentes para cada traço.

Lembrete: Uma razão 2:1 é produzida, em geral, por um gene letal.

$$Yy \times Yy$$

$\downarrow$

YY	1/4 morre
Yy	1/2 amarelo, se torna 2/3
yy	1/4 cinza, se torna 1/3

Agora, vamos examinar a herança do tipo de pelo. Dois camundongos com pelo liso são cruzados, produzindo 1/2 + 1/4 = 2/4 + 1/4 = 3/4 camundongos com pelo liso e 1/6 + 1/12 = 2/12 + 1/12 = 3/12 = 1/4 com pelo encrespado. Aprendemos no Capítulo 3 que uma razão 3:1 é produzida por um cruzamento entre dois indivíduos heterozigotos para um único alelo dominante:

Dica: O Quadro 3.5 fornece as razões fenotípicas para cruzamentos genéticos simples.

$$Ss \times Ss$$

$\downarrow$

SS	1/4 liso	3/4 liso
Ss	1/2 liso	
ss	1/4 encrespado	

Agora podemos combinar ambos os *loci* e indicar os genótipos para todos os camundongos no cruzamento:

P *Yy Ss* × *Yy Ss*

Amarelo, liso Amarelo, liso

↓

Fenótipo	Genótipo	Probabilidade em cada *locus*	Probabilidade combinada
Amarelo, liso	*Yy S_*	$2/3 \times 3/4$	$= 6/12 = 1/2$
Amarelo, encrespado	*Yy ss*	$2/3 \times 1/4$	$= 2/12 = 1/6$
Cinza, liso	*yy S_*	$1/3 \times 3/4$	$= 3/12 = 1/4$
Cinza, encrespado	*yy ss*	$1/3 \times 1/4$	$= 1/12$

b. Poderíamos fazer vários cruzamentos diferentes para testar nossa hipótese de que amarelo é um recessivo letal e liso é dominante sobre encrespado. Por exemplo, um cruzamento entre dois camundongos amarelos sempre produz 2/3 de amarelo e 1/3 de cinza nos descendentes e um cruzamento entre dois camundongos cinza deve produzir apenas descendentes cinza. Um cruzamento entre dois camundongos com pelo encrespado produz apenas descendentes com esse pelo.

Problema 2

Em alguns carneiros, a existência de chifres é produzida por um alelo autossômico que é dominante nos machos e recessivo nas fêmeas. Uma fêmea com chifres é cruzada com um macho sem chifres. Uma das fêmeas de F_1 é cruzada com um macho sem chifre. Qual proporção de machos e fêmeas entre os descendentes deste cruzamento terá chifres?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
A proporção de machos e fêmeas dos descendentes com chifres.

Quais informações são fornecidas para solucionar o problema?

- Os símbolos para os alelos de chifre e sem chifre
- Chifres são produzidos por um alelo autossômico que é dominante nos machos e recessivo nas fêmeas
- Uma fêmea com chifres é cruzada com um macho sem chifres. Uma fêmea de F_1 é cruzada com um macho sem chifre para produzir descendentes.

Para a solução deste problema, revise:
Características influenciadas e características limitadas pelo sexo, na Seção 5.3.

Etapas para a solução

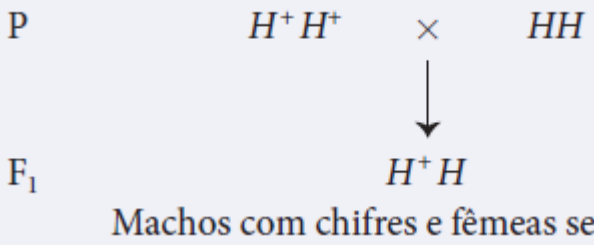
A existência de chifres nesses carneiros é um exemplo de uma característica influenciada pelo sexo. Como os fenótipos associados aos genótipos são diferentes para os dois sexos, vamos começar o problema escrevendo os genótipos e os fenótipos para cada sexo. Vamos usar *H* para representar o alelo que codifica os chifres e H^+ para o alelo que codifica sem chifres. Nos machos, o alelo para chifres é dominante sobre o alelo para sem chifres, o que significa que os machos homozigotos *(HH*) e heterozigotos (H^+H) para este gene têm chifres. Apenas os machos homozigotos para o alelo recessivo (H^+H^+) não têm chifres. Nas fêmeas, o alelo para chifres é recessivo, o que significa que apenas as fêmeas homozigotas para este alelo *(HH*) têm chifres, as fêmeas heterozigotas (H^+H) e homozigotas (H^+H^+) para o alelo sem chifres não têm a característica. O quadro a seguir resume os genótipos e fenótipos associados:

Dica: Escreva os genótipos e o fenótipo associado para cada sexo.

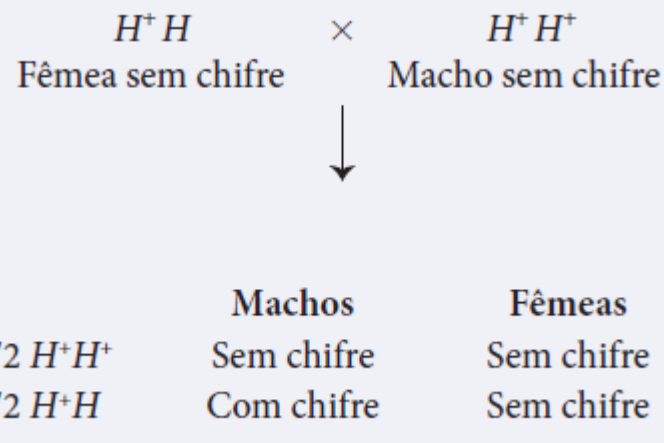
Lembrete: Quando um traço é dominante, o homozigoto e heterozigoto o expressam nos seus fenótipos.

Genótipo	Fenótipo do macho	Fenótipo da fêmea
HH	Com chifre	Com chifre
HH$^+$	Com chifre	Sem chifre
H$^+$*H*$^+$	Sem chifre	Sem chifre

No problema, uma fêmea com chifres é cruzada com um macho sem chifres. A partir do quadro fornecido, observamos que uma fêmea com chifre tem de ser homozigota para o alelo de chifre (*HH*) e um macho sem chifre tem de ser homozigoto para o alelo sem chifre (*H*$^+$*H*$^+$), então todos os descendentes F_1 serão heterozigotos, os machos F_1 terão chifres e as fêmeas F_1 não terão, como é demonstrado no diagrama a seguir:



Uma fêmea heterozigota sem chifre F_1 (*H*$^+$*H*) é cruzada com um macho sem chifre (*H*$^+$*H*$^+$):



Portanto, 1/2 dos descendentes machos terá chifre, mas nenhuma fêmea terá.

Questões de compreensão

Seção 5.1

1. Qual é a diferença entre dominância incompleta e codominância?
2. O que é penetrância incompleta e o que provoca este fenômeno?

Seção 5.2

3. O que é interação gênica? Qual é a diferença entre um gene epistático e um gene hipostático?
4. O que é um gene epistático recessivo?
5. O que é teste de complementação e para que serve?

Seção 5.3

6. Quais características são exibidas por um traço de herança citoplasmática?
7. O que é *imprinting* genômico?
8. Qual é a diferença entre efeito genético materno e *imprinting* genômico?
9. Qual é a diferença entre um gene influenciado pelo sexo e um gene que exibe *imprinting* genômico?

Seção 5.4

10. Quais características você espera ver em um traço que exibe antecipação?

Seção 5.5

11. O que são características contínuas e de que modo elas surgem?

Questões e problemas aplicados

Seções 5.1 a 5.4

12. Combine cada um dos itens a seguir com sua definição correta (itens *a* a *i*).

___ fenocópia
___ efeito genético materno
___ pleiotropia
___ *imprinting* genômico
___ traço influenciado pelo sexo
___ traço limitado pelo sexo
___ antecipação
___ penetrância
___ traço poligênico

a. A porcentagem de organismos com um genótipo específico que expressa o fenótipo associado.
b. Um traço determinado por um gene autossômico que é mais facilmente expresso em um sexo.
c. Um traço determinado por um gene autossômico que é expresso em apenas um sexo.
d. Um traço que é determinado por um efeito ambiental e tem o mesmo fenótipo como um traço geneticamente determinado.
e. Um traço determinado por genes em vários *loci*.
f. A expressão de um traço é afetada pelo sexo do genitor que transmite o gene para os descendentes.
g. O traço aparece mais cedo ou de forma mais grave nas gerações seguintes.
h. Um gene que afeta mais de um fenótipo.
i. O genótipo da genitora determina o fenótipo dos descendentes.

Seção 5.1

*13. Os cavalos palomino têm pelo amarelo dourado, os cavalos castanhos têm pelo marrom e os cavalos cremello têm pelo que é quase branco. Vários cruzamentos entre os três tipos diferentes de cavalos produziram os seguintes descendentes:

Cruzamento	Descendentes
Palomino × palomino	13 palominos, 6 castanhos, 5 cremellos
Castanho × castanho	16 castanhos
Cremello × cremello	13 cremellos
Palomino × castanho	8 palominos, 9 castanhos
Palomino × cremello	11 palominos, 11 cremellos
Castanho × cremello	23 palominos

a. Explique a herança dos fenótipos palomino, castanho e cremello nos cavalos.
b. Indique símbolos para os alelos que determinam estes fenótipos e liste os genótipos de todos os genitores e descendentes apresentados na tabela anterior.

14. Os alelos L^M e L^N no *locus* do tipo sanguíneo MN exibem codominância. Indique os genótipos e fenótipos e suas razões nos descendentes dos seguintes cruzamentos.
a. $L^ML^M \times L^ML^N$.
b. $L^NL^N \times L^NL^N$.
c. $L^ML^N \times L^ML^N$.

d. $L^M L^N \times L^N L^N$.
e. $L^M L^M \times L^N L^N$.

*15. Suponha que os lóbulos longos de orelha nos seres humanos sejam um traço autossômico dominante que exibe 30% de penetrância. Uma pessoa heterozigota para os lóbulos longos de orelha acasala com uma pessoa homozigota para lóbulos normais de orelha. Qual é a probabilidade de sua primeira criança ter lóbulos longos de orelha?

16. O pé torto é uma das anomalias esqueléticas congênitas mais comuns, com uma incidência mundial em torno de 1 em cada 1.000 nascimentos. Acredita-se que fatores genéticos e não genéticos sejam responsáveis por esta anomalia. C. A. Gurnett *et al*. (*American Journal of Human Genetics*. 2008, 83:616-622) identificou uma família na qual o pé torto era um traço autossômico dominante com penetrância reduzida. Eles descobriram uma mutação no gene *PITXI*, responsável pelo pé torto. Por meio do teste de DNA, eles determinaram que 11 pessoas da família tinham a mutação *PITXI*, mas apenas 8 delas tinham pé torto. Qual é a penetrância da mutação *PITXI* nesta família?

*17. Quando um hamster chinês com manchas brancas é cruzado com outro hamster sem manchas, aproximadamente 1/2 dos descendentes tem manchas brancas e 1/2, não. Quando dois hamsters com manchas brancas são cruzados, 2/3 dos descendentes têm manchas brancas e 1/3, não.
a. Qual é a base genética da mancha branca nos hamsters chineses?
b. Como você faria para produzir hamsters chineses que tenham filhotes com manchas brancas?

18. No início do século 20, Lucien Cuénot estudou os dados da base genética do pelo amarelo nos camundongos (discutido na p. 97). Ele realizou vários cruzamentos entre dois camundongos amarelos e obteve o que ele pensou que fosse uma razão 3:1 de camundongos amarelos para cinza nos descendentes. O quadro a seguir mostra os resultados verdadeiros de Cuénot, junto com os resultados de uma série muito maior de cruzamentos realizada por Castle e Little (W E. Castle, C. C. Little. 1910. *Science* 32:868-870).

Descendentes resultantes dos cruzamentos de camundongos amarelo × amarelo

Investigadores	Descendentes amarelos	Descendentes não amarelos	Total de descendentes
Cuénot	263	100	363
Castle e Little	800	435	1.235
Ambos combinados	1.063	535	1.598

a. Usando um teste qui-quadrado, determine se os resultados de Cuénot são muito diferentes da razão 3:1 que ele pensou ter observado. Eles são muito diferentes de uma razão 2:1?
b. Determine se os resultados de Castle e Little são muito diferentes de uma razão 3:1. Eles são diferentes de uma razão 2:1?
c. Combine os resultados de Castle e Cuénot e determine se eles são muito diferentes de uma razão 3:1 e uma razão 2:1.
d. Apresente uma explicação para as diferentes razões que Cuénot e Castle tiveram.

*19. No milheto, a cor é determinada por três alelos em um único *locus*: Rp^1 (vermelho), Rp^2 (púrpura), e *rp* (verde). O vermelho é dominante sobre púrpura e verde, e púrpura é dominante sobre verde ($Rp^1 > Rp^2 > rp$). Indique os fenótipos e as razões esperados dos descendentes produzidos pelos seguintes cruzamentos.
a. $Rp^1/Rp^2 \times Rp^1/rp$.
b. $Rp^1/rp \times Rp^2/rp$.
c. $Rp^1/Rp^2 \times Rp^1/Rp^2$.
d. $Rp^2/rp \times rp/rp$.
e. $rp/rp \times Rp^1/Rp^2$.

20. Se existem cinco alelos em um *locus*, quantos genótipos podem existir neste *locus*? Quantos tipos diferentes de homozigotos existem para este alelo? Quantos genótipos e homozigotos podem existir com oito alelos em um *locus*?

21. Os perus têm uma plumagem preta, bronze ou azul-bronze. Examine os resultados dos seguintes cruzamentos:

Genitores	Descendentes
Cruzamento 1: preto e bronze	todos pretos

Cruzamento 2: preto e preto	3/4 preto, 1/4 bronze
Cruzamento 3: preto-bronze e preto-bronze	todos preto-bronze
Cruzamento 4: preto e bronze	1/2 preto, 1/4 bronze, 1/4 preto-bronze
Cruzamento 5: bronze e preto-bronze	1/2 bronze, 1/2 preto-bronze
Cruzamento 6: bronze e bronze	3/4 bronze, 1/4 preto-bronze

Você acha que essas diferenças na plumagem surgem por causa de uma dominância incompleta entre dois alelos em um único *locus*? Se sim, justifique sua conclusão ao indicar símbolos para cada alelo e fornecer os genótipos para todos os perus nos cruzamentos. Se sua resposta for não, forneça uma explicação alternativa e indique os genótipos para todos os perus nos cruzamentos.

22. Nos coelhos, uma série alélica ajuda a determinar a cor do pelo: *C* (sem mancha), c^{ch} (chinchila, cinza), c^{h} (himalaio, branco com extremidades pretas) e *c* (albino, todo branco). O alelo *C* é dominante sobre todos os outros, c^{ch} é dominante sobre c^{h} e *c*, c^{h} é dominante sobre *c*, e *c* é recessivo para todos os outros. Esta hierarquia de dominância pode ser resumida como $C > c^{ch} > c^{h} > c$. Os coelhos na lista a seguir foram cruzados e produziram os descendentes apresentados. Indique os genótipos dos genitores para cada cruzamento:

	Fenótipos dos genitores	**Fenótipos dos descendentes**
a.	Sem mancha × albino	1/2 sem mancha, 1/2 albino
b.	Himalaio × albino	1/2 himalaio, 1/2 albino
c.	Sem mancha × albino	1/2 sem mancha, 1/2 chinchila
d.	Sem mancha × himalaio	1/2 sem mancha, 1/4 himalaio, 1/4 albino
e.	Sem mancha × sem mancha	3/4 sem mancha, 1/4 albino

23. Neste capítulo, consideramos a ação de paternidade de Joan Barry contra Charlie Chaplin e como, com base nos tipos sanguíneos, Chaplin não poderia ser o pai da criança.
    a. Quais tipos sanguíneos são possíveis para o pai da filha de Barry?
    b. Se Chaplin tivesse um destes tipos sanguíneos, isto provaria que ele é o pai da criança?

*24. Uma mulher tem tipo sanguíneo AM. Ela tem uma criança com tipo sanguíneo AB MN. Quais dos tipos sanguíneos abaixo *não* seria o tipo sanguíneo do pai da criança? Explique seu raciocínio.

George	O	N
Tom	AB	MN
Bill	B	MN
Claude	A	N
Henry	AB	M

Seção 5.2

*25. Nas galinhas, o formato da crista é determinado por alelos em dois *loci* (*R, r* e *P, p*). Uma crista em noz é produzida quando pelo menos um alelo dominante *R* está presente em um *locus* e pelo menos um alelo dominante *P* está presente em um segundo *locus* (genótipo *R_P_*). Uma crista em rosa é produzida quando pelo menos um alelo dominante *R* está presente no primeiro *locus* e dois alelos recessivos estão presentes em um segundo *locus* (genótipo *R_pp*). Uma crista em ervilha é produzida quando existem dois alelos recessivos no primeiro *locus* e pelo menos um alelo dominante no segundo (genótipo *rr P_*). Se houver dois alelos recessivos no primeiro e no segundo *loci* (*rr pp*), uma crista simples é produzida. Os descendentes com quais tipos de cristas, e em que proporções, surgiram a partir dos cruzamentos abaixo?
    a. *RR PP* × *rr pp*
    b. *Rr Pp* × *rr pp*
    c. *Rr Pp* × *Rr Pp*
    d. *Rr pp* × *Rr pp*
    e. *Rr pp* × *rr Pp*
    f. *Rr pp* × *rr pp*

*26. Tatuo Aida investigou a base genética da variação da cor do peixe Medaka (*Aplocheilus latipes*), um pequeno peixe encontrado no Japão (T. Aida. 1921. *Genetics* 6:554-573), e descobriu que genes em dois *loci* (*B, b* e *R r*) determinam a cor do peixe: peixes com um alelo dominante em ambos *loci* (*B_R_*) são marrons, peixes com um alelo dominante apenas

no *locus B* (*B_ rr*) são azuis, peixes com apenas um alelo dominante no *locus R* (*bb R_*) são vermelhos, e peixes com alelos recessivos em ambos *loci* (*bb rr*) são brancos. Aida cruzou um peixe marrom homozigoto com um peixe branco homozigoto. Ele fez o retrocruzamento de F_1 com o genitor branco homozigoto e obteve 228 peixes marrons, 230 peixes azuis, 237 peixes vermelhos e 222 peixes brancos.

a. Indique os genótipos dos descendentes do retrocruzamento.
b. Use um teste qui-quadrado para comparar os números observados dos descendentes do retrocruzamento com os números esperados. Que conclusão você pode tirar a partir dos resultados do qui-quadrado?
c. Quais resultados você esperaria de um cruzamento entre um peixe vermelho homozigoto e um peixe branco?
d. Quais resultados você esperaria se cruzasse um peixe vermelho homozigoto com um peixe azul homozigoto e, então, retrocruzasse F_1 com o genitor vermelho homozigoto?

27. Uma variedade da papoula (*Papaver somniferum* L.) com folhas denteadas foi cruzada com uma variedade com folhas normais. Todas as plantas F_1 tinham folhas denteadas. Duas plantas de F_1 são intercruzadas para produzir F_2. De F_2, 249 tinham folhas denteadas e 16 tinham folhas normais. Indique os genótipos para todas as plantas das gerações de P, F_1 e F_2. Explique como as folhas denteadas eram determinadas na papoula.

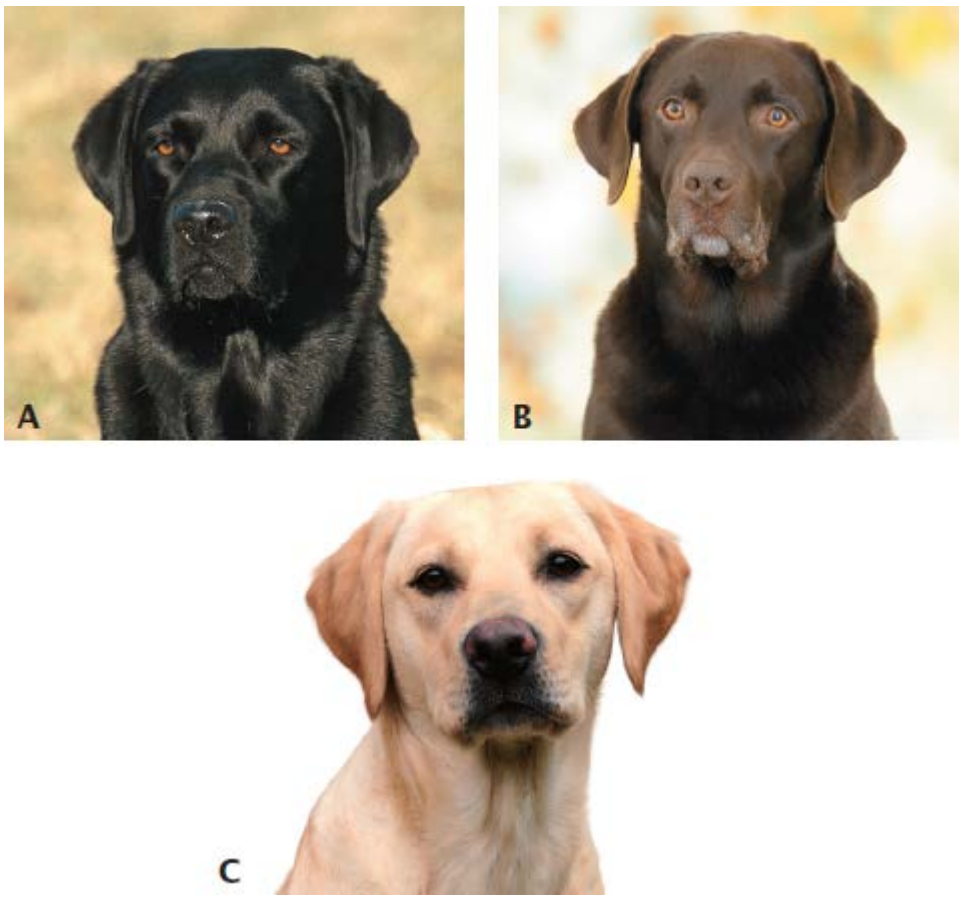
ANÁLISE DE DADOS

28. E. W. Lindstrom cruzou dois milhos com mudas verdes e obteve os seguintes descendentes: 3.583 mudas verdes, 853 mudas branco-esverdeadas e 260 mudas amarelas (E. W. Lindstrom 1921. *Genetics* 6:91-110).

a. Forneça os genótipos para os descendentes verdes, branco-esverdeados e amarelos.
b. Explique como a cor é determinada nestas mudas.
c. Há epistasia entre os genes que determinam a cor nas mudas de milho? Se existe, qual gene é epistático e qual é hipostático?

*29. Um criador de cães gostava dos labradores retriever amarelo e marrom. Em uma tentativa para produzir filhotes amarelos e marrons, ele cruzou um labrador-macho amarelo com um labrador-fêmea marrom. No entanto, todos os filhotes deste cruzamento eram pretos (ver p. 101 sobre uma discussão da base genética da cor do pelo nos labradores retrievers).

a. Explique este resultado.
b. Como o criador conseguiria obter labradores amarelos e marrons?



Cor do pelo no labrador retriever: **A.** Preto. **B.** Marrom. **C.** Amarelo. (*Partes A* e *B*: Juniors Bildarchiv/Alamy. *Parte C*: C. Byatt-Norman/Shutterstock.)

30. Quando um labrador-fêmea amarela foi cruzada com um macho marrom, metade dos filhotes era marrom e a outra metade era amarela. A mesma fêmea, quando cruzada com um macho marrom diferente, teve apenas descendentes marrons. Explique esses resultados.

*31. Uma abóbora que produz o fruto com formato de disco é cruzada com uma abóbora que produz o fruto longo. Todos os descendentes F_1 têm formato de disco. Quando os descendentes F_1 são intercruzados, os descendentes F_2 são produzidos na seguinte razão: 9/16 frutos em formato de disco, 6/16 frutos esféricos e 1/16 fruto longo. Indique os genótipos dos descendentes F_2.

32. Algumas ervilhas-de-cheiro têm flores púrpura e outras têm flores brancas. Uma variedade homozigota de ervilha com flores púrpura é cruzada com uma variedade homozigota com flores brancas. Todas as plantas F_1 têm flores púrpura. Quando as plantas F_1 são autofertilizadas, F_2 aparece em uma razão de 9/16 púrpura e 7/16 brancas.
    a. Forneça os genótipos para as flores púrpura e brancas nestes cruzamentos.
    b. Desenhe uma via bioquímica hipotética para explicar a produção das flores púrpura e brancas nas ervilhas-de-cheiro.

*33. Consulte as pp. 106-107 sobre uma discussão de como a cor e o padrão do pelo são determinados nos cães.
    a. Porque os setters irlandeses são avermelhados?
    b. Um poodle cruzando com qualquer outra raça pode produzir filhotes manchados? Por que sim ou não?
    c. Se um são-bernardo for cruzado com um dobermann, qual será a cor do pelo dos descendentes: sem mancha, amarela, sela ou bicolor?
    d. Se um rottweiler é cruzado com um labrador retriever, qual será a cor do pelo dos descendentes: sem mancha, amarela, sela ou bicolor?

Seção 5.3

34. A puberdade precoce limitada ao homem resulta de um alelo autossômico, raro, limitado ao sexo (*P*) que é dominante sobre o alelo para puberdade normal (*p*) e é expresso apenas nos homens. Bill sofre de puberdade precoce, mas seu irmão Jack e sua irmã Bete entraram na puberdade no tempo certo, entre 10 e 14 anos. Embora os pais de Bill tenham entrado na puberdade na época certa, dois tios maternos (irmãos de sua mãe) sofreram puberdade precoce. Todos os avôs de Bill tiveram puberdade normal. Indique os genótipos mais prováveis para todos os parentes mencionados nesta família.

*35. Em algumas cabras, a existência de chifres é produzida por um alelo autossômico que é dominante nos machos e recessivo nas fêmeas. Uma fêmea com chifres é cruzada com um macho sem chifres. Os descendentes F_1 são intercruzados para produzir F_2. Qual proporção de fêmeas F_2 que terá chifre?

36. Nas cabras, a existência de barba é produzida por um alelo autossômico que é dominante nos machos e recessivo nas fêmeas. Usaremos o símbolo B^b para o alelo de barba e B^+ para o alelo sem barba. Outro alelo autossômico, que se separa de forma independente, produz um pelo preto (*W*) dominante sobre o alelo para o pelo branco (*w*). Indique os fenótipos e suas proporções esperadas para os seguintes cruzamentos.
    a. B^+B^b *Ww* macho × B^+B^b *Ww* fêmea.
    b. B^+B^b *Ww* macho × B^+B^b *ww* fêmea.
    c. B^+B^+ *Ww* macho × B^bB^b *Ww* fêmea.
    d. B^+B^b *Ww* macho × B^bB^b *ww* fêmea.

37. As penas em franjas nas galinhas são um traço autossômico recessivo limitado ao sexo para os machos. Liste todos os possíveis genótipos para as galinhas apresentadas em:
    a. Figura 5.13 A.
    b. Figura 5.13 B.

38. J. K. Breitenbecher (1921. *Genetics* 6:65-86) investigou a base genética da variação de cor do gorgulho do feijão de quatro manchas (*Bruchus quadrimaculatus*). Os gorgulhos são vermelho, preto, branco ou bege. Breitenbecher descobriu que quatro alelos (*R*, R^b, R^w e *r*) em um único *locus* determinam a cor. Os alelos exibem uma hierarquia de dominância, com o vermelho (*R*) dominante sobre todos os outros alelos, preto (R^b) dominante sobre branco (R^w) e bege (*r*), branco dominante sobre bege e bege recessivo para todos os outros ($R > R^b > R^w > r$). Os seguintes genótipos codificam cada uma das cores:

RR, RR^b, RR^w, *Rr*	Vermelha
R^bR^b, R^bR^w, R^br	Escura
R^wR^w, R^wr	Branca
rr	Bege

A variação de cor nesta espécie é limitada ao sexo para as fêmeas: os machos carreiam genes de cor, mas são sempre bege, independentemente do seu genótipo. Para cada um dos cruzamentos realizados por Breitenbecher, indique todos os possíveis genótipos dos genitores.

	Genitores	Descendentes
a.	Bege ♀ × bege ♂	78 vermelhos ♀, 70 brancos ♀, 184 bege ♂
b.	Preto ♀ × bege ♂	151 vermelhos ♀, 49 pretos ♀, 61 bege ♀, 249 bege ♂
c.	Branco ♀ × bege ♂	32 vermelhos ♀, 31 bege ♂
b.	Preto ♀ × bege ♂	3.586 pretos ♀, 1.282 bege ♀, 4.791 bege ♂
e.	Branco ♀ × bege ♂	594 brancos ♀, 189 bege ♀, 862 bege ♂
f.	Preto ♀ × bege ♂	88 pretos ♀, 88 bege ♀, 186 bege ♂
g.	Bege ♀ × bege ♂	47 brancos ♀, 51 bege ♀, 100 bege ♂
h.	Vermelho ♀ × bege ♂	1.932 vermelhos ♀, 592 bege ♀, 2.587 bege ♂
i.	Branco ♀ × bege ♂	13 vermelhos ♀, 6 brancos ♀, 5 bege ♀, 19 bege ♂
j.	Vermelho ♀ × bege ♂	190 vermelhos ♀, 196 pretos ♀, 311 bege ♂
k.	Preto ♀ × branco ♀	1.412 pretos, 502 brancos, 1.766 bege ♂

*39. A enroscadura do casco do caracol *Lymnaea peregra* (discutido na introdução do capítulo) surge a partir de um efeito genético materno. Um alelo autossômico para o casco para a direita, ou destro (s^+), é dominante sobre o alelo para o casco para a esquerda, ou canhoto (s). Um caracol de criação chamado Marta é canhoto e se reproduz apenas como fêmea (os caracóis são hermafroditas). Indique quais afirmações a seguir são verdadeiras e quais são falsas. Justifique sua resposta para cada caso.
   a. O genótipo da Marta *tem de* ser *ss*.
   b. O genótipo da Marta *não pode* ser s^+s^+.
   c. Todos os descendentes produzidos por Marta *têm de* ser canhotos.
   d. Pelo menos alguns descendentes produzidos por Marta *têm de* ser canhotos.
   e. A genitora de Marta *tem de* ser canhota.
   f. Todos os irmãos de Marta *têm de* ser canhotos.

40. Se os caracóis destros F_2 com genótipo s^+s na Figura 5.17 sofrem autofertilização, quais fenótipos e proporções são esperados nos descendentes?

41. A hipospadia, um defeito de nascença nos homens nos quais a uretra se abre no corpo em vez de na ponta do pênis, resulta de um gene autossômico dominante em algumas famílias. As fêmeas que carreiam o gene não apresentam efeitos. Este defeito congênito é um exemplo de (a) traço ligado ao X, (b) traço ligado ao Y, (c) um traço limitado pelo sexo, (d) um traço influenciado pelo sexo ou (e) efeito genético materno? Explique sua resposta.

42. Nos unicórnios, dois *loci* autossômicos interagem para determinar o tipo de rabo. Um *locus* controla a existência de rabo; o alelo para rabo (*T*) é dominante sobre o alelo para sem rabo (*t*). Se um unicórnio tem um rabo, então os alelos em um segundo *locus* determinam se o rabo é enrolado ou liso. O fazendeiro Baldrigde tem dois unicórnios com rabos enrolados; quando eles foram cruzados, 1/2 dos descendentes tinha rabo enrolado, 1/4 tinha rabos lisos e 1/4 não tinha rabo. Forneça os genótipos dos genitores e descendentes do cruzamento do fazendeiro Baldridge. Explique como ele obteve uma razão fenotípica 2:1:1 neste cruzamento.

43. Em 1983, um criador de carneiros em Oklahoma observou no seu rebanho um carneiro que tinha massa muscular aumentada nos quartos traseiros. Muitos dos descendentes desse carneiro tinham o mesmo traço, que se tornou conhecido como o mutante *callipyge* (*callipyge* em grego significa “lindo traseiro”). A mutação que gerava o fenótipo *callipyge* acabou sendo encontrada no cromossomo 18 do carneiro.

    Quando os descendentes machos *callipyge* do carneiro mutante original foram cruzados com fêmeas normais, eles produziram os seguintes descendentes: 1/4 macho *callipyge*, 1/4 fêmea *callipyge*, 1/4 macho normal e 1/4 fêmea normal. Quando a fêmea *callipyge* descendente do carneiro mutante original foi cruzada com machos normais, todos seus descendentes eram normais. A análise dos cromossomos desses descendentes mostrou que metade deles recebeu um cromossomo 18 com o alelo que codificava a característica *callipyge* da sua genitora. Proponha uma explicação para a herança do alelo para *callipyge*. Como você poderia testar sua explicação?

Seção 5.5

44. Qual das afirmações a seguir é um exemplo de uma fenocópia? Explique seu raciocínio.
   a. A fenilcetonúria é o resultado de uma mutação recessiva que provoca pele clara assim como déficit intelectual.

b. A altura do ser humano é influenciada por genes em muitos *loci* diferentes.
c. As plantas anãs e folhas matizadas nos tomates são causadas por genes diferentes que estão ligados.
d. As asas vestigiais na *Drosophila* são produzidas por uma mutação recessiva. Este traço também é produzido por alta temperatura durante o desenvolvimento.
e. A inteligência nos seres humanos é influenciada por fatores genéticos e ambientais.

*45. As orelhas longas em alguns cães são um traço autossômico dominante. Dois cães cruzam e produzem uma ninhada na qual 75% dos filhotes têm orelhas longas. Dos cães com orelhas longas nesta ninhada, 1/3 eram fenocópias. Quais são os genótipos mais prováveis para os genitores desta ninhada?

46. A mosca com asas vestigiais apresentada no canto inferior à esquerda da Figura 5.19 é cruzada com a mosca com asas normais mostrada no canto superior à direita da figura. Se os descendentes forem criados a 31°C, qual porcentagem terá asas vestigiais?

Questões desafiadoras

Seção 5.1

47. Os pombos são objeto de estudos genéticos há muito tempo. De fato, Charles Darwin criou pombos na esperança de descobrir os princípios da hereditariedade, mas não teve sucesso. Várias investigações genéticas no início do século 20 pesquisaram a base hereditária da variação de cor nestes pássaros. W. R. Horlancher estava interessado na base genética do *kiteness*, um padrão de cor que é uma mistura de pontos vermelhos e pretos das penas. Ele fez os seguintes cruzamentos para investigar a relação genética do *kiteness* para a cor da pena preta e vermelha (W. R. Horlancher. 1930. *Genetics* 15:312-346).

Cruzamento	Descendentes
Kitey × *kitey*	16 *kitey*, 5 pretos, 3 vermelhos
Kitey × preto	6 *kitey*, 7 pretos
Vermelho × *kitey*	18 vermelhos, 9 *kitey*, 6 pretos

a. Com base nestes resultados, proponha uma hipótese para explicar a herança das cores de pena *kitey*, preta e vermelha nos pombos. (Dica: Considere que dois *loci* estão envolvidos e ocorre algum tipo de epistasia.)
b. Para cada um dos cruzamentos anteriores, teste sua hipótese usando um teste qui-quadrado.

Seção 5.3

48. Suponha que você está cuidando de uma colônia de camundongos em um instituto de pesquisa genética e um dia você descobre um camundongo com orelhas deformadas. Você cruza esse camundongo e descobre que o traço é herdado. Ambos os camundongos macho e fêmea têm orelhas deformadas, mas quando você cruza um macho com orelhas deformadas com uma fêmea de orelhas normais, seus resultados são diferentes dos resultados obtidos quando você cruza uma fêmea de orelhas deformadas com um macho com orelhas normais: os cruzamentos recíprocos geram resultados diferentes. Descreva como você determinaria se este traço é resultado de um gene ligado ao sexo, gene influenciado pelo sexo, um efeito genético materno, um gene herdado do citoplasma ou *imprinting* genômico. Quais cruzamentos você faria e quais resultados seriam esperados com estes diferentes tipos de herança?

6

Análise de Heredograma, Aplicações e Exame Genético

O mistério das impressões digitais desaparecidas

Em 2007, uma suíça de 29 anos tentou entrar nos EUA. Embora sua aparência estivesse de acordo com a foto no seu passaporte, ela não conseguiu fazer a verificação das impressões digitais, não porque suas impressões fossem as de um terrorista, mas porque nenhum de seus dedos tinha qualquer impressão. Ela ficou retida por horas, pois os fiscais da imigração ficaram confusos em decidir o que fazer com uma pessoa sem impressões digitais.

As impressões digitais são um dos nossos traços únicos e permanentes. Ninguém – mesmo entre gêmeos idênticos – compartilha as impressões digitais com outra pessoa. As impressões digitais são chamadas tecnicamente de cristas epidérmicas ou padrões dermatoglíficos, encontrados nos nossos dedos das mãos, dos pés, nas palmas das mãos e solas dos pés. As cristas epidérmicas surgem bem antes do nascimento – elas estão totalmente formadas 17 semanas após a concepção e são permanentes. Pesquisas mostram que os padrões das impressões digitais são claramente influenciados pela hereditariedade, mas também existem fatores aleatórios. Um dos primeiros cientistas a estudar as impressões digitais foi Francis Galton, um primo de Charles Darwin. No final do século 19, Galton estabeleceu que duas pessoas não têm as mesmas impressões digitais e mostrou que as impressões dos parentes são mais semelhantes que as de pessoas sem parentesco.

A completa ausência das impressões, como ocorreu com a passageira suíça no aeroporto, é uma condição muito rara, conhecida como adermatoglifia (ADG). Apelidada de “doença do atraso da imigração” em razão do embaraço que ela causa quando pessoas com essa condição tentam atravessar fronteiras, a ADG foi documentada apenas em algumas pessoas de quatro famílias no mundo. Na ADG, as impressões digitais estão ausentes no nascimento e nunca se desenvolvem. O distúrbio não produz efeitos prejudiciais.

Em 2011, geneticistas em Israel e na Suíça solucionaram o mistério das impressões digitais ausentes nas pessoas com ADG. Janna Nousbeck e seus colegas examinaram essa condição em uma grande família suíça na qual alguns membros tinham impressões digitais normais e outros membros, não (**Figura 6.1**). Nessa família, a ADG exibe as características principais de um traço autossômico dominante: ela ocorre igualmente entre homens e mulheres, não pula gerações, e todas as pessoas com a condição têm um parente com a mesma condição. Os pesquisadores coletaram amostras de sangue de membros da família que não tinham impressões digitais e de parentes com impressões normais. Eles extraíram o DNA do sangue e fizeram genotipagem dos familiares para 6.000 polimorfismos de nucleotídio único (SNPs), sequências de DNA que apresentam variação em um único nucleotídio. Ao comparar a existência de SNPs nos membros da família com e sem impressões digitais, eles foram capazes de determinar que o gene para ADG estava localizado em um intervalo específico no braço longo do cromossomo 4. Um dos genes na região é *SMARCAD1*, que codifica uma forma curta de uma proteína encontrada exclusivamente na pele. O sequenciamento do gene revelou que os membros da família com ADG tinham uma mutação não encontrada nos membros com impressões digitais. A mutação provoca recomposição ou *splicing* anormal no RNA transcrito do gene, de modo que esse RNA é menos estável. Ainda não se esclareceu como a estabilidade reduzida desse RNA causa ADG, mas os cientistas esperam que a identificação desse gene leve a uma melhor compreensão de como as impressões digitais se desenvolvem.

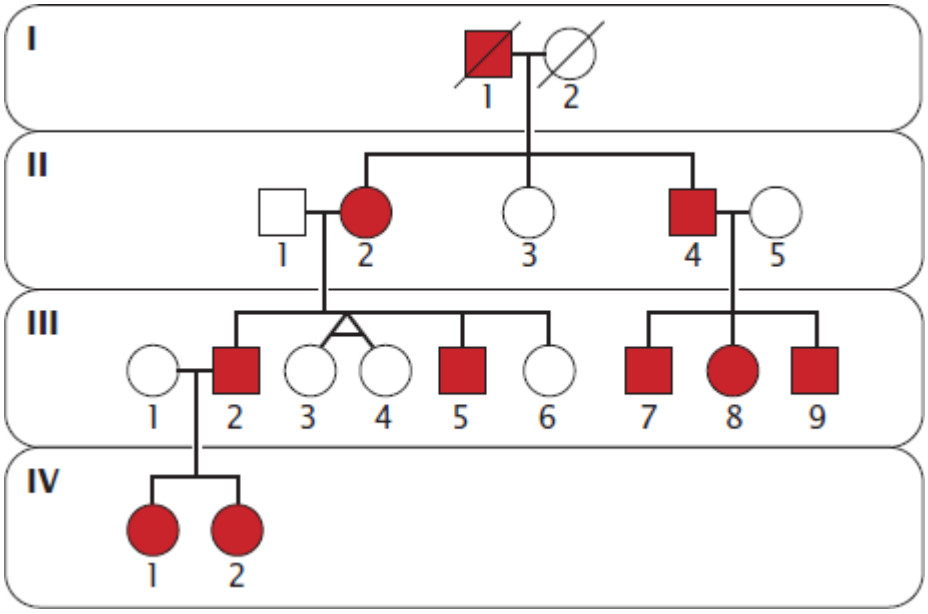


Figura 6.1 Heredograma de família suíça com adermatoglifia (ausência de impressões digitais). Os quadrados representam os homens, os círculos representam as mulheres. Os quadrados e círculos coloridos são as pessoas com adermatoglifia.

A ausência de impressões digitais é um dos vários traços e doenças nos seres humanos que atualmente são o foco de intensa pesquisa genética. Neste capítulo, vamos estudar as características genéticas e examinar três técnicas importantes usadas pelos geneticistas para investiga-las: heredograma, estudos com gêmeos e estudos de adoção. No final do capítulo, veremos como as informações obtidas com essas técnicas podem ser usadas no aconselhamento genético e no diagnóstico pré-natal.

Tenha em mente, ao ler este capítulo, que muitas características importantes são influenciadas pelos genes e pelo meio ambiente, e é sempre difícil separar tais fatores nos seres humanos. Os estudos com gêmeos e pessoas adotadas são projetados para distinguir os efeitos dos genes e do meio ambiente, mas eles se baseiam em suposições que podem ser difíceis de serem estabelecidas para algumas características humanas, em particular as características comportamentais. Portanto, é prudente interpretar os resultados desses estudos com cautela.

6.1 O estudo da genética nos seres humanos é limitado por características especiais da biologia e da cultura humanas

Os seres humanos são os melhores e os piores organismos para o estudo genético. Por um lado, sabemos mais sobre anatomia, fisiologia e bioquímica humana que sobre a maioria dos outros organismos, então muitos traços bem caracterizados estão disponíveis para estudo. As famílias guardam registros detalhados de seus membros que se estendem por muitas gerações. Além disso, várias doenças humanas importantes têm um componente genético; então o incentivo para compreender a herança humana é imenso. Por outro lado, o estudo das características genéticas humanas tem grandes obstáculos.

Primeiramente, acasalamentos controlados não são possíveis. Os geneticistas fazem cruzamentos específicos com outros organismos para testar suas hipóteses sobre herança. Como já observamos, o cruzamento-teste, por exemplo, fornece uma forma conveniente para determinar se um organismo com um traço dominante é homozigoto ou heterozigoto. Infelizmente (para os geneticistas pelo menos), os acasalamentos entre os seres humanos são determinados pelo romance, as expectativas das famílias ou, em alguns casos, acidente, em vez das exigências de um geneticista.

Outro obstáculo é o de que os seres humanos têm um tempo de geração longo. A idade fértil humana não é alcançada antes dos 10 a 14 anos após o nascimento e a maioria das pessoas não se reproduz antes dos 18 anos ou mais, formando um tempo de geração, em geral, de 20 anos. Esse tempo longo de geração significa que, mesmo se os geneticistas pudessem controlar os acasalamentos entre os seres humanos, eles teriam que esperar uma média de 40 anos apenas para observar os descendentes F_2. Por outro lado, o tempo de geração da *Drosophila* é de 2 semanas; nas bactérias, apenas 20 min.

Finalmente, a família dos seres humanos é, em geral, pequena. A observação até de razões genéticas simples que aprendemos no Capítulo 3 exigiria um número substancial de descendentes em cada família. Quando os pais geram apenas duas crianças, é impossível detectar uma razão 3: 1. Mesmo em uma família muito grande com 10 a 15 crianças, não seria possível identificar uma razão di-híbrida 9: 3: 3: 1.

Embora essas limitações tornem os estudos genéticos nos seres humanos mais complexos, é muito importante compreender a hereditariedade humana. Portanto, os geneticistas foram forçados a desenvolver técnicas adequadas apenas para a biologia e a cultura humanas. Resolva o Problema 18

Conceitos

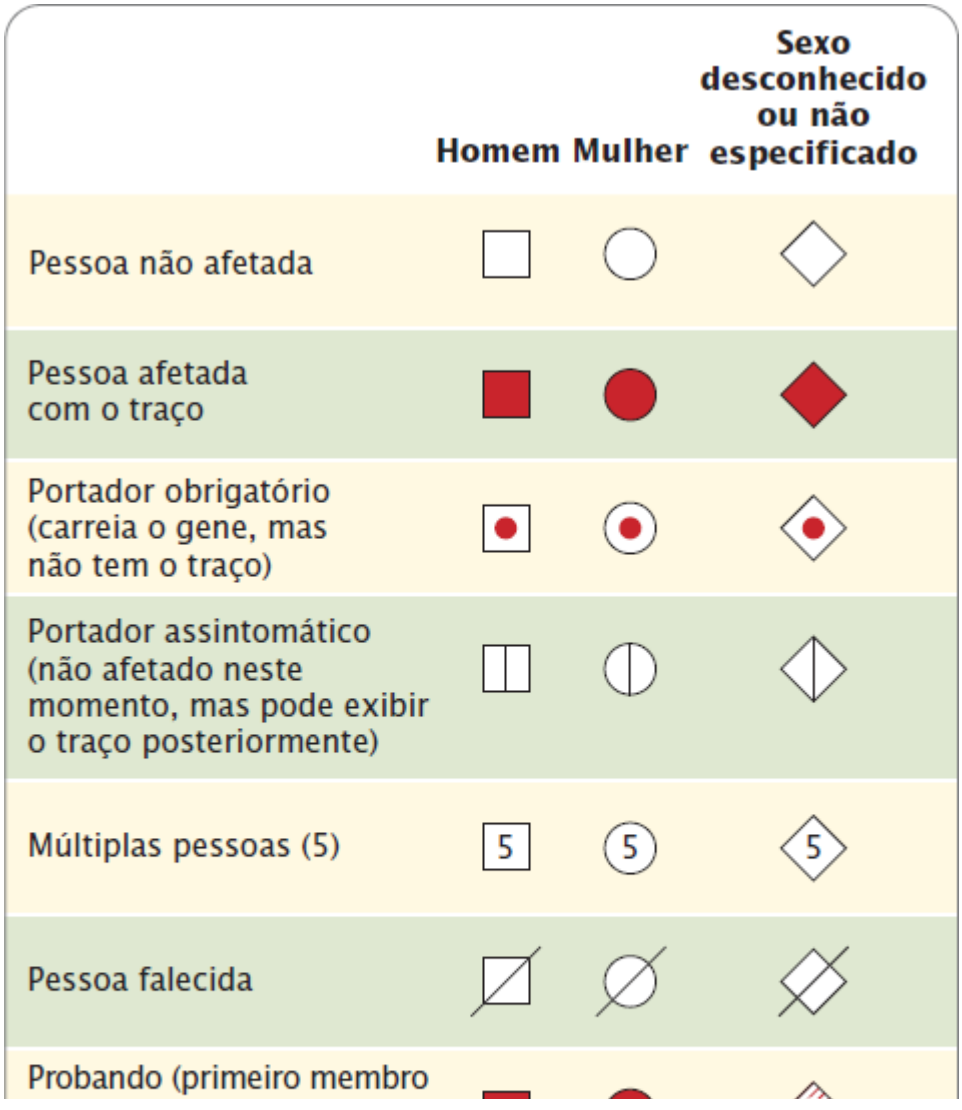
Embora os princípios de hereditariedade sejam os mesmos nos seres humanos e em outros organismos, o estudo da hereditariedade humana é limitado pela incapacidade de se controlar os cruzamentos genéticos, pelo longo tempo de geração e pelo pequeno número de descendentes.

6.2 Os geneticistas usam frequentemente os heredogramas para estudar a hereditariedade das características nos seres humanos

Uma técnica importante usada pelos geneticistas para estudar a herança humana é a análise dos heredogramas. Um **heredograma** é uma representação pictórica da história familiar, basicamente uma árvore genealógica que esboça a herança de uma ou mais características. Quando uma característica ou doença específica é observada em uma pessoa, o geneticista estuda a família dela em um heredograma.

Símbolos dos heredogramas

Os símbolos mais usados nos heredogramas estão resumidos na **Figura 6.2**. Os homens são representados por quadrados, e as mulheres, por círculos. Uma linha horizontal desenhada entre dois símbolos representando um homem e uma mulher indica um acasalamento; as crianças são conectadas aos seus pais por linhas verticais que se estendem abaixo de seus pais. O heredograma apresentado na **Figura 6.3 A** ilustra uma família com a síndrome de Waardenburg, um tipo de surdez autossômica dominante que pode ser acompanhado por pele clara, topete branco e problemas visuais (**Figura 6.3 B**). As pessoas que exibem esse traço estão representadas por círculos e quadrados cheios no heredograma na Figura 6.3 A. Os membros não afetados estão representados por círculos e quadrados vazios. A pessoa a partir da qual o heredograma é iniciado chama-se **probando** e, em geral, é indicada por uma seta (IV-2 na Figura 6.3 A).



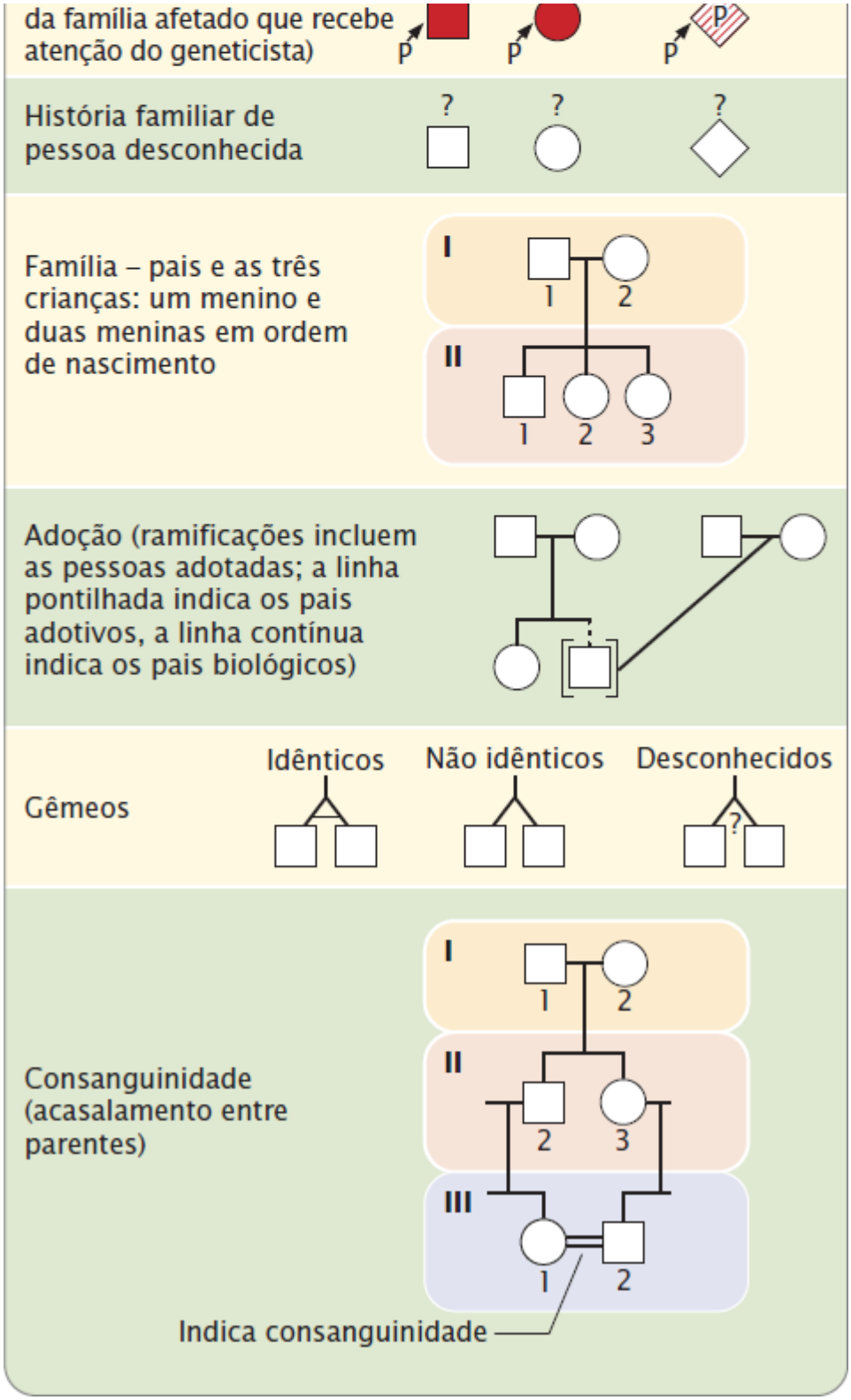


Figura 6.2 **Símbolos padrão são usados nos heredogramas.**

Vamos examinar com cuidado a Figura 6.3 e considerar algumas características adicionais de um heredograma. Cada geração no heredograma é identificada por um número romano; dentro de cada geração, os membros são indicados por números arábicos, e as crianças em cada família estão listadas em ordem de nascimento da esquerda para a direita. A pessoa II-4, um homem com a síndrome de Waardenburg, copulou com II-5, uma mulher não afetada, e eles tiveram cinco filhos. O mais velho é III-8, um homem com a síndrome de Waardenburg e a mais jovem é III-14, uma mulher não afetada. Resolva o Problema 19a

Análise dos heredogramas

O número limitado de descendentes na maioria das famílias humanas significa que geralmente é impossível discriminar claras razões mendelianas em um único heredograma. A análise de heredograma requer um pouco de investigação, com base na identificação dos padrões associados aos diferentes modos de herança. Por exemplo, os traços autossômicos dominantes devem aparecer com a mesma frequência em ambos os sexos e não devem pular geração, desde que o traço seja totalmente penetrante e

não influenciado pelo sexo (ver *Capítulo 5*).

Alguns padrões podem excluir a possibilidade de um modo específico de herança. Por exemplo, um filho herda seu cromossomo X de sua mãe. Se observarmos que um traço é transmitido do pai para o filho, podemos excluir a possibilidade de herança ligada ao X. Os traços discutidos nas seções a seguir serão considerados totalmente penetrantes e raros.

Traços autossômicos recessivos

Normalmente, os traços autossômicos recessivos surgem com a mesma frequência em ambos os sexos (exceto se a penetrância for diferente em homens e mulheres) e aparecem apenas quando uma pessoa herda dois alelos para o traço, um de cada genitor. Se o traço não é comum, a maioria dos pais dos descendentes afetados é heterozigota e não afetada e, assim, temos a impressão de que o traço pula gerações (**Figura 6.4**). Frequentemente, um alelo recessivo pode ser passado para várias gerações sem que o traço apareça em um heredograma. Independentemente de ambos os pais serem heterozigotos, espera-se que aproximadamente 25% dos descendentes expressem o traço, mas essa razão pode não ser evidente, exceto se a família for grande. No caso raro de ambos os pais serem afetados por um traço autossômico recessivo, todos os descendentes serão afetados.

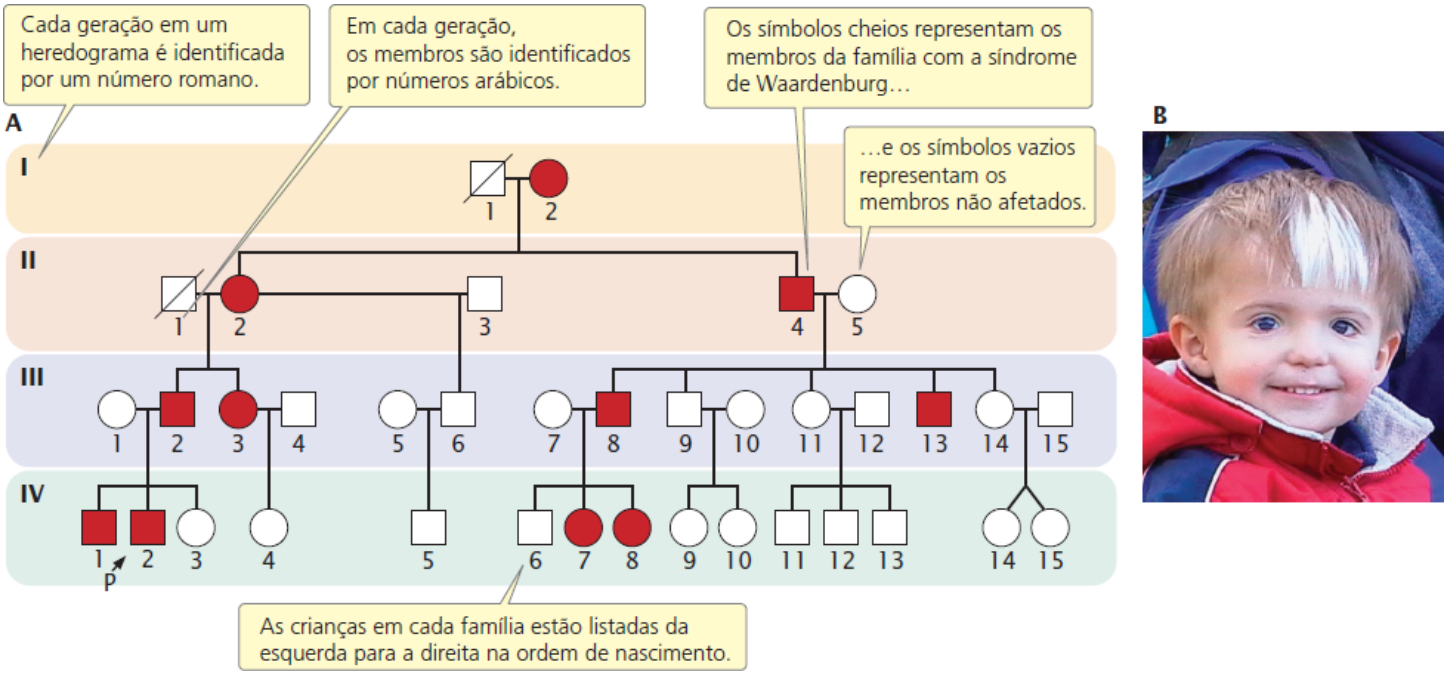


Figura 6.3 A síndrome de Waardenburg é (A) herdada como um traço autossômico dominante e (B) caracterizada por surdez, pele clara, problemas visuais e topete branco. O probando (*P*) é a pessoa a partir da qual esse heredograma é iniciado. (Cortesia de Guy Rowland.)

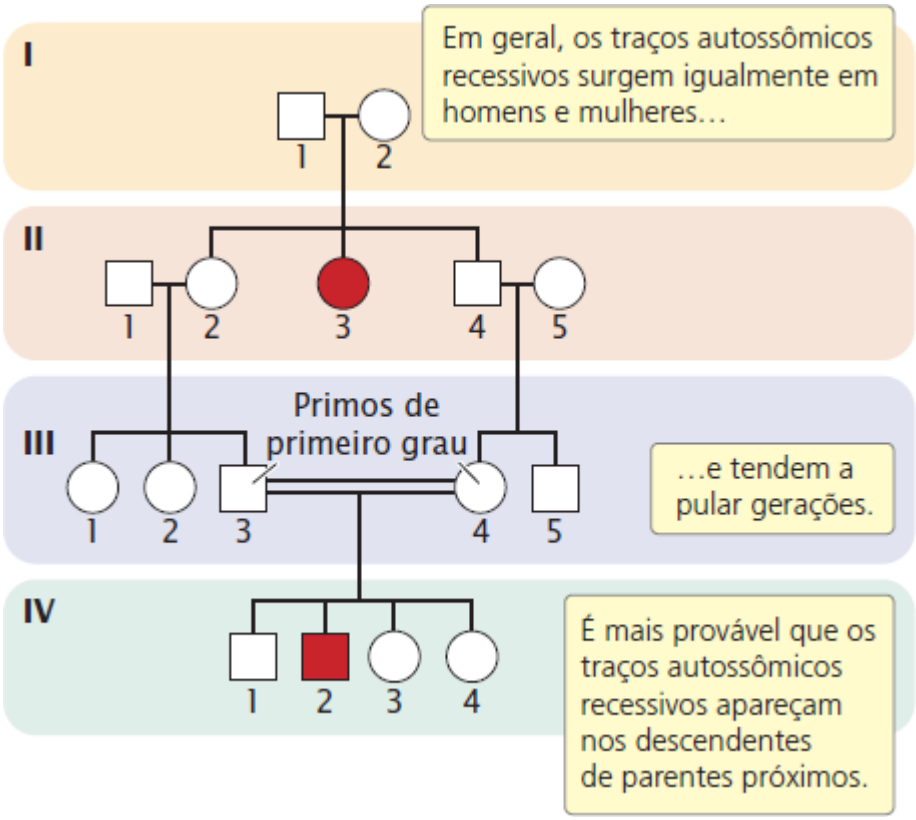


Figura 6.4 Os traços autossômicos recessivos normalmente aparecem com a mesma frequência em ambos os sexos e tendem a pular gerações.

Quando um traço recessivo é raro, as pessoas fora da família são, em geral, homozigotas para o alelo normal. Assim, quando uma pessoa afetada acasala com uma pessoa de fora da família (*aa* × *AA*), em geral, nenhuma das crianças apresentará o traço, embora todas sejam portadoras (ou seja, heterozigotas). É mais provável que um traço recessivo apareça em um heredograma quando duas pessoas da mesma família acasalam, porque existe uma chance maior de ambos os pais carrearem o mesmo alelo recessivo. O acasalamento de parentes próximos é chamado de **consanguinidade**. No heredograma apresentado na Figura 6.4, as pessoas III-3 e III-4 são primos em primeiro grau e ambas são heterozigotas para o alelo recessivo; quando dois heterozigotos acasalam, é esperado que 25% das crianças tenham o traço recessivo.

Conceitos

Os traços autossômicos recessivos surgem com a mesma frequência em homens e mulheres. É comum que as crianças afetadas sejam filhos de pais não afetados que são portadores do gene para o traço, e esse traço tende a pular gerações. Os traços recessivos aparecem mais frequentemente nos descendentes de acasalamentos consanguíneos.

Checagem dos conceitos 1

Os traços autossômicos recessivos surgem nos heredogramas com acasalamentos consanguíneos, visto que eles:

a. Tendem a pular gerações.
b. Aparecem apenas quando ambos os pais carreiam uma cópia para o traço, que é mais provável quando os pais são parentes.
c. Em geral surgem nos filhos de pessoas não afetadas.
d. Surgem igualmente em homens e mulheres.

Várias doenças metabólicas nos seres humanos são herdadas como traços autossômicos recessivos. Um exemplo é a doença de Tay-Sachs. As crianças com a doença de Tay-Sachs parecem saudáveis quando nascem, mas se tornam apáticas e fracas com 6 meses de vida. Seu quadro físico e neurológico piora gradualmente, levando à cegueira, surdez e, consequentemente, morte entre 2 e 3 anos de idade. A doença é o resultado do acúmulo de um lipídio chamado gangliosídeo G_{M2} no cérebro. Um componente normal das células do cérebro, o gangliosídeo G_{M2} é degradado por uma enzima chamada hexosaminidase A,

ausente nas crianças com a doença de Tay-Sachs. Esse gangliosídeo em excesso se acumula no cérebro, causando edema e, por último, manifestações neurológicas. As pessoas heterozigotas têm apenas uma cópia normal do alelo que codifica a hexosaminidase A e produzem apenas metade da quantidade normal dessa enzima. Entretanto, essa quantidade é suficiente para garantir que o gangliosídeo G_{M2} seja degradado normalmente, e os heterozigotos, normais.

Traços autossômicos dominantes

Os traços autossômicos dominantes surgem em ambos os sexos com a mesma frequência, e ambos são capazes de transmitir esses traços para seus descendentes. Cada pessoa com um traço dominante necessariamente herdou o alelo de pelo menos um genitor; dessa forma, os traços autossômicos dominantes não pulam gerações (**Figura 6.5**). As exceções para essa regra surgem quando as pessoas adquirem o traço como o resultado de uma nova mutação ou quando o traço tem penetrância reduzida.

Se um alelo autossômico dominante for raro, a maioria das pessoas que apresenta esse traço é heterozigota. Quando um dos genitores é heterozigoto e afetado e o outro não é afetado, aproximadamente 50% dos descendentes serão afetados. Se ambos os pais têm o traço e são heterozigotos, aproximadamente 75% dos descendentes serão afetados. As pessoas não afetadas não transmitem o traço para seus descendentes, desde que ele seja totalmente penetrante. Na Figura 6.5, observamos que nenhum dos descendentes de II-4 (que não é afetado) tem o traço.

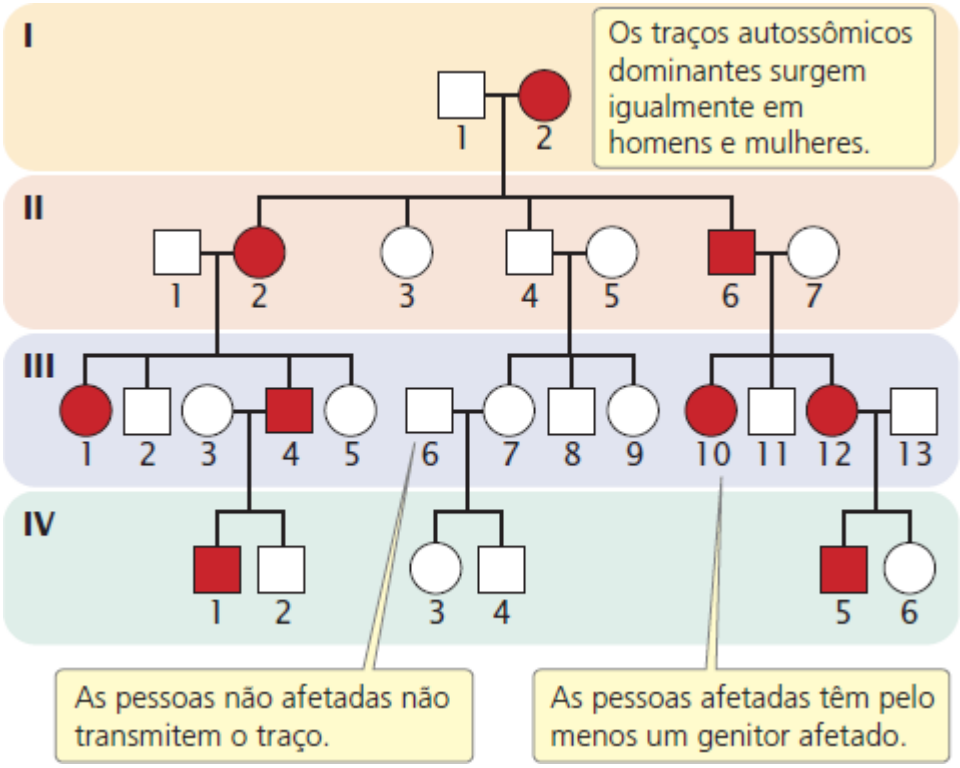


Figura 6.5 **Os traços autossômicos dominantes aparecem com a mesma frequência em ambos os sexos e não pulam gerações.**

Conceitos

Os traços autossômicos dominantes surgem em ambos os sexos com a mesma frequência. Uma pessoa afetada tem um genitor afetado (exceto se a pessoa tiver novas mutações) e o traço não pula gerações. As pessoas que não são afetadas não transmitem o traço.

Checagem dos conceitos 2

Quando você pode observar um traço autossômico dominante pulando gerações?

Um traço geralmente considerado autossômico dominante é a hipercolesterolemia familiar, uma doença hereditária na qual o colesterol sanguíneo está muito elevado devido a um defeito no transporte de colesterol. O colesterol é transportado no corpo em pequenas partículas solúveis chamadas de lipoproteínas (**Figura 6.6**). Uma lipoproteína importante no transporte do colesterol é a lipoproteína de baixa densidade (LDL). Quando uma molécula de LDL alcança uma célula, ela se liga a um receptor para LDL, que desloca a LDL através da membrana celular para o citoplasma, onde é degradada e seu colesterol é

liberado para ser usado pela célula.

A hipercolesterolemia familiar ocorre quando existe um defeito no gene que normalmente codifica o receptor para LDL. Em geral, a doença é considerada um distúrbio autossômico dominante porque os heterozigotos têm déficit de receptores para LDL e níveis sanguíneos de colesterol elevados, levando a aumento do risco de doença da artéria coronária. As pessoas heterozigotas para hipercolesterolemia familiar têm níveis sanguíneos de LDL que são duas vezes o normal e sofrem infarto do miocárdio aos 35 anos.

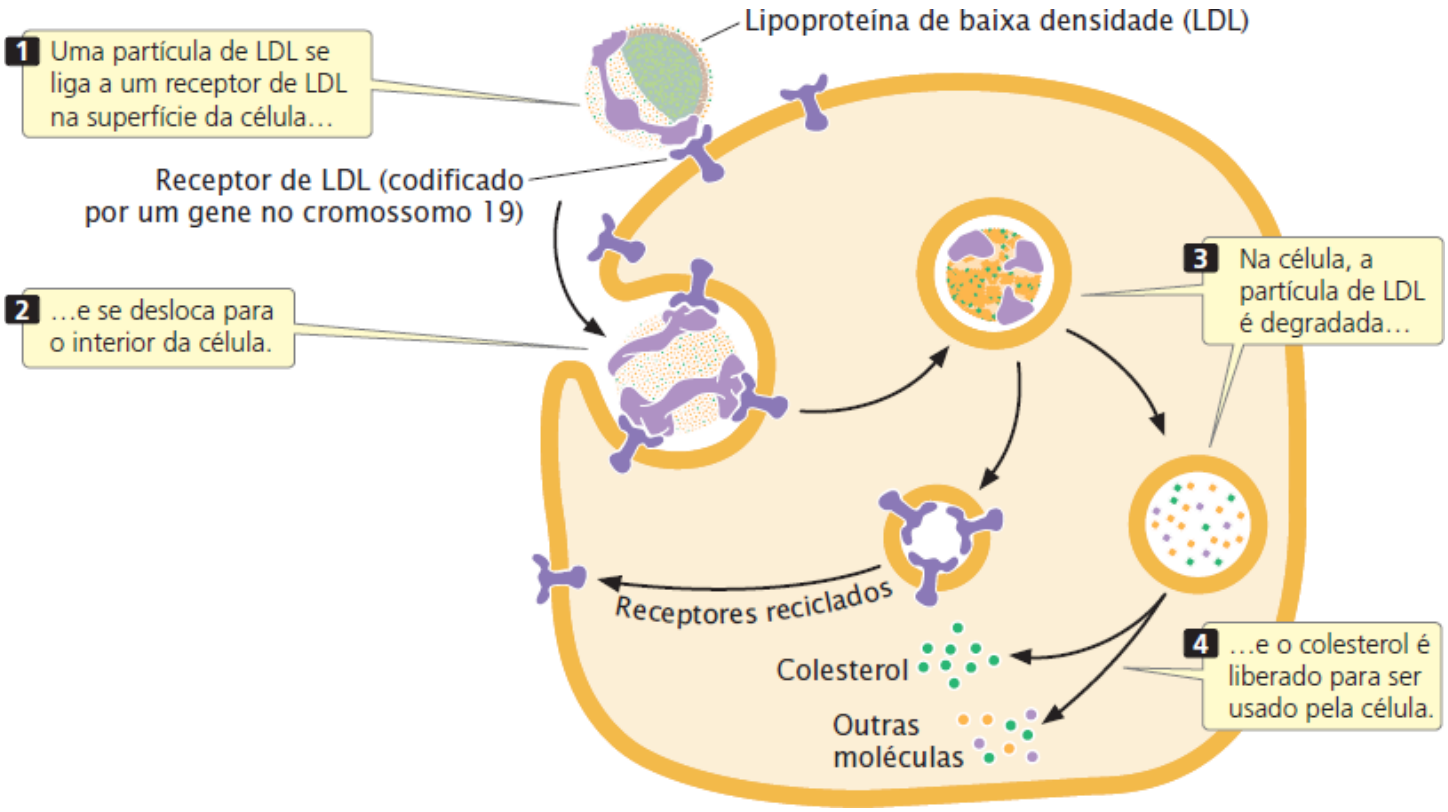


Figura 6.6 As partículas de lipoproteína de baixa densidade (LDL) transportam o colesterol. O receptor de LDL desloca a LDL da corrente sanguínea através da membrana celular, para dentro do citoplasma.

Em casos muito raros, uma pessoa herda dois alelos defeituosos para o receptor de LDL. Essas pessoas não produzem nenhum receptor funcional para LDL, seus níveis de colesterol no sangue são seis vezes acima do normal e elas podem infartar antes dos 2 anos e quase inevitavelmente até os 20. Como os homozigotos são afetados de forma mais grave de que os heterozigotos, a hipercolesterolemia familiar é considerada um traço com dominância incompleta. Entretanto, é raro encontrar homozigotos, e a forma comum da doença nos heterozigotos aparece como um traço dominante simples na maioria dos heredogramas.

Traços recessivos ligados ao X

Os traços recessivos ligados ao X têm um padrão diferente de herança (**Figura 6.7**). Primeiro, eles surgem mais frequentemente nos homens que nas mulheres, porque eles precisam herdar apenas uma única cópia do alelo para apresentar o traço, enquanto as mulheres precisam herdar duas, uma de cada genitor, para serem afetadas. Segundo, como um homem herda seu cromossomo X da mãe, é comum os homens afetados nascerem de mulheres não afetadas portadoras do alelo para o traço. Como o traço é transmitido da mulher não afetada para o homem afetado e, então, para a mulher não afetada, ele tende a pular gerações (ver Figura 6.7). Quando uma mulher é heterozigota, aproximadamente 50% de seus filhos serão afetados e 50% de suas filhas serão portadoras não afetadas. Por exemplo, sabemos que as mulheres I-2, II-2 e III-7 na Figura 6.7 são portadoras porque transmitiram o traço para aproximadamente 50% de seus filhos.

Uma terceira característica importante dos traços recessivos ligados ao X é que eles não são transmitidos do pai para o filho, porque um filho herda o seu cromossomo Y do pai, não o X. Na Figura 6.7, não existe caso no qual pai e filho sejam afetados. Todas as filhas de um homem afetado, entretanto, serão portadoras (se sua mãe for homozigota para o alelo normal). Quando uma mulher apresenta um traço recessivo ligado ao X, ela tem de ser homozigota para esse traço e todos seus filhos também o apresentarão.

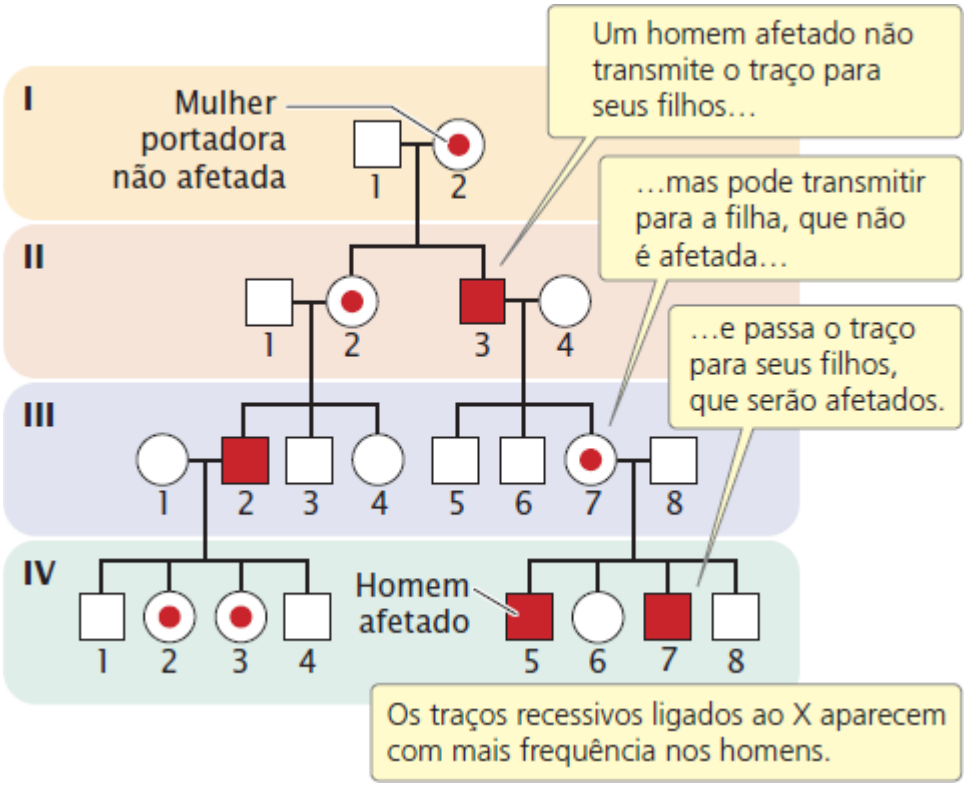


Figura 6.7 Os traços recessivos ligados ao X aparecem com maior frequência nos homens que nas mulheres e não são transmitidos do pai para o filho.

Conceitos

Raros traços recessivos ligados ao X aparecem com maior frequência nos homens que nas mulheres e não são transmitidos do pai para o filho. Os homens afetados em geral são filhos de mulheres não afetadas que são portadoras do gene para o traço; assim, os traços recessivos ligados ao X tendem a pular gerações.

Checagem dos conceitos 3

Como você pode diferenciar um traço autossômico recessivo com maior penetrância nos homens de um traço recessivo ligado ao X?

Um exemplo de traço recessivo ligado ao X nos seres humanos é a hemofilia A, também chamada de hemofilia clássica. A hemofilia é o resultado da ausência de uma proteína necessária para a coagulação sanguínea. O processo complexo de coagulação sanguínea é uma cascata de reações que inclui mais de 13 fatores diferentes. Por isso, existem vários tipos de distúrbios de coagulação, cada um causado por um pequeno defeito em uma etapa diferente da via de coagulação.

A hemofilia A é o resultado de um fator VIII ausente ou anormal, uma das proteínas na cascata de coagulação. O gene para o fator VIII está localizado na extremidade do braço longo do cromossomo X; então a hemofilia A é um distúrbio recessivo ligado ao X. As pessoas com hemofilia A sangram em excesso e até pequenos cortes e contusões são potencialmente fatais. Ocorre sangramento espontâneo nas articulações como cotovelos, joelhos e tornozelos, produzindo dor, edema e erosão dos ossos. Felizmente, o sangramento nas pessoas com esse tipo de hemofilia pode ser controlado com a administração de doses concentradas do fator VIII. A herança da hemofilia está ilustrada na família da Rainha Vitória da Inglaterra (**Figura 6.8**).

Traços dominantes ligados ao X

Os traços dominantes ligados ao X aparecem em homens e mulheres, embora sejam mais frequentes nas mulheres. Cada pessoa com um traço dominante ligado ao X tem obrigatoriamente um genitor afetado (exceto se a pessoa tiver uma nova mutação ou se o traço for de penetrância reduzida). Os traços dominantes ligados ao X não pulam gerações (**Figura 6.9**); os homens afetados transmitem o traço para todas as suas filhas e para nenhum de seus filhos, como é observado nas crianças de I-1 na Figura 6.9. Por outro lado, as mulheres afetadas (se forem heterozigotas) transmitem o traço para cerca de 1/2 de seus filhos e 1/2 de suas filhas, como é observado nas crianças de III-6 no heredograma. Como ocorre com os traços recessivos ligados ao X, um homem herda um traço dominante ligado ao X apenas de sua mãe; o traço não é transmitido do pai para o filho. Esse fato é que diferencia a herança dominante ligada ao X da herança autossômica dominante, na qual um homem pode herdar o traço de

seu pai. Uma mulher, por outro lado, herda um cromossomo X de sua mãe e de seu pai, então as mulheres podem receber um traço dominante ligado ao X de qualquer um de seus genitores.

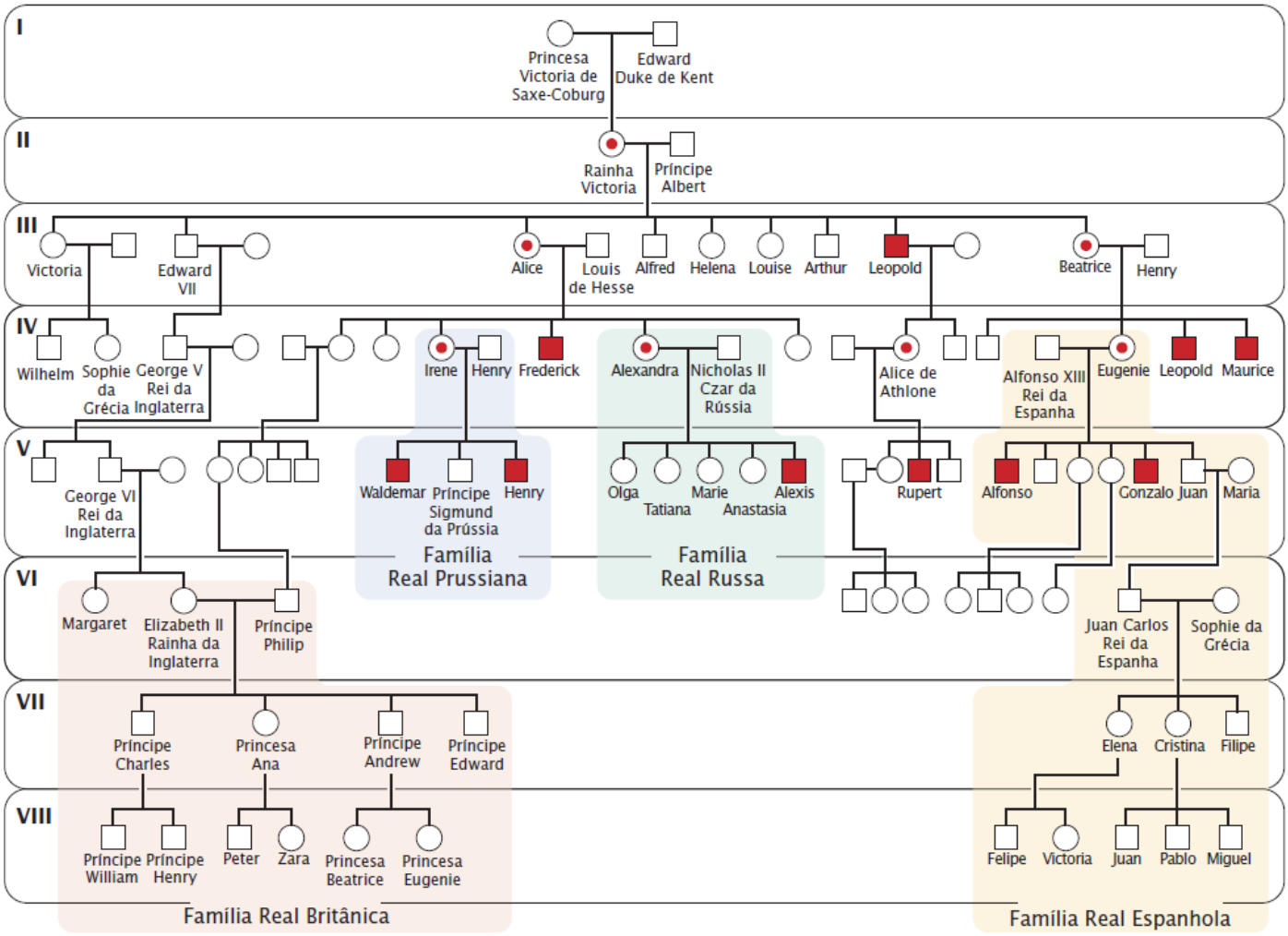


Figura 6.8 A hemofilia clássica é herdada como um traço recessivo ligado ao X. Este heredograma representa a hemofilia nas famílias reais da Europa.

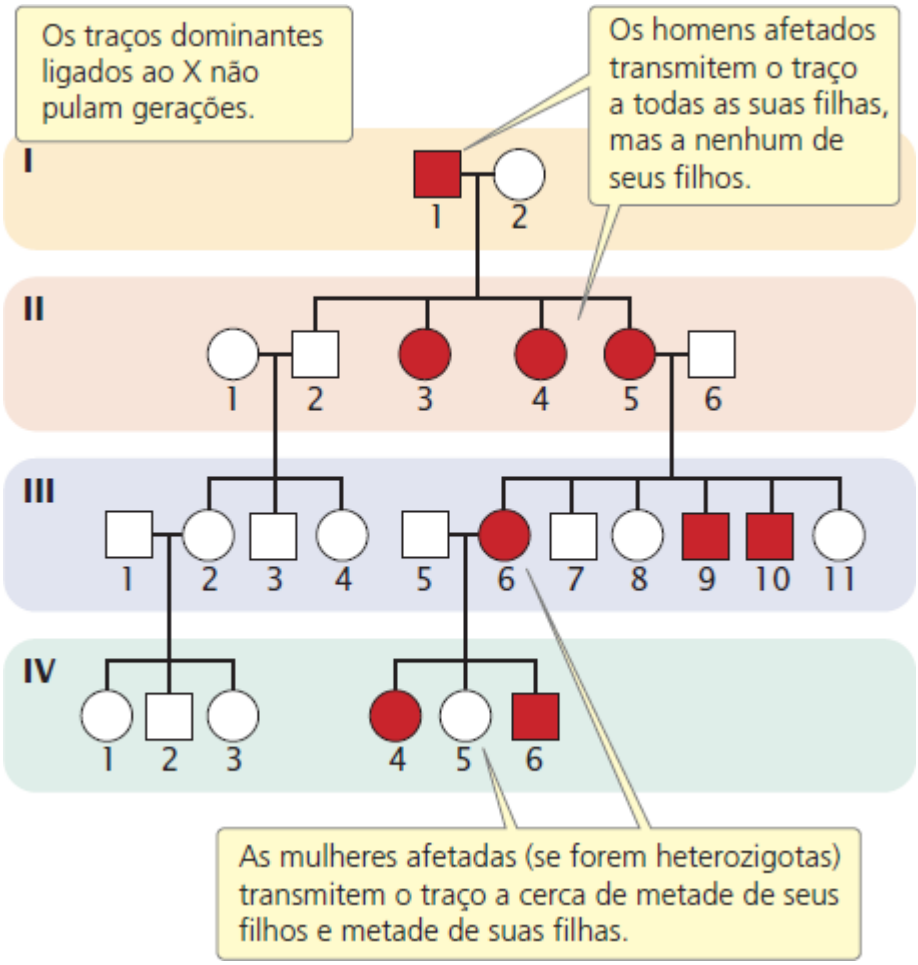


Figura 6.9 Os traços dominantes ligados ao X afetam homens e mulheres. Um homem afetado tem obrigatoriamente a mãe afetada.

Conceitos

Os traços dominantes ligados ao X afetam homens e mulheres. Os homens afetados têm necessariamente mães afetadas (exceto se os homens tiverem uma nova mutação) e eles transmitem o traço para todas as suas filhas.

Checagem dos conceitos 4

Um homem afetado com um traço dominante ligado ao X terá que proporção de descendentes afetada pelo traço?

a. 1/2 de filhos e 1/2 de filhas.
b. Todos os filhos e nenhuma das filhas.
c. Todas as filhas e nenhum dos filhos.
d. 3/4 das filhas e 1/4 dos filhos.

Um exemplo de traço dominante ligado ao X nos seres humanos é a hipofosfatemia, também conhecida como raquitismo familiar resistente a vitamina D. As pessoas com esse traço têm características que lembram superficialmente o raquitismo: deformidades ósseas, coluna vertebral e articulações rígidas, joelho varo e leve deficiência de crescimento. Esse distúrbio, entretanto, é resistente ao tratamento com a vitamina D, que normalmente cura o raquitismo. A hipofosfatemia ligada ao X resulta do transporte defeituoso do fosfato, em especial nas células dos rins. Pessoas com esse distúrbio excretam muito fosfato na urina, resultando em níveis baixos no sangue e depósito reduzido de minerais nos ossos. O distúrbio é tratado com altas doses de calcitriol (uma forma hormonal ativa da vitamina D) e fosfato. Como ocorre com os traços dominantes ligados ao X, os homens com hipofosfatemia são, com frequência, afetados de forma mais grave que as mulheres.

Traços ligados ao Y

Os traços ligados ao Y exibem um padrão específico, de fácil identificação, de herança. Apenas os homens são afetados e o

traço é transmitido de pai para filho. Se um homem for afetado, todos os seus descendentes masculinos também serão afetados, como é o caso de I-1, II-4, II-6, III-6 e III-10 do heredograma na **Figura 6.10**. Os traços ligados ao Y não pulam gerações. Como foi mencionado no Capítulo 4, existe pouca informação genética no cromossomo Y humano. A masculinidade é um dos poucos traços nos seres humanos que parece ligado ao Y. Uma vez que cada homem tem apenas um único cromossomo Y, existe apenas uma cópia de cada alelo ligado ao Y. Assim, os traços ligados ao Y não são dominantes nem recessivos.

Conceitos

Os traços ligados ao Y são encontrados apenas nos homens e são transmitidos de um pai para todos seus filhos.

Checagem dos conceitos 5

Quais características de um heredograma diferenciariam um traço ligado ao Y de um traço raro, autossômico dominante e limitado pelo sexo nos homens?

As principais características dos traços autossômicos recessivos, autossômicos dominantes, recessivos ligados ao X, dominantes ligados ao X e ligados ao Y estão resumidas no **Quadro 6.1**. Resolva o Problema 22

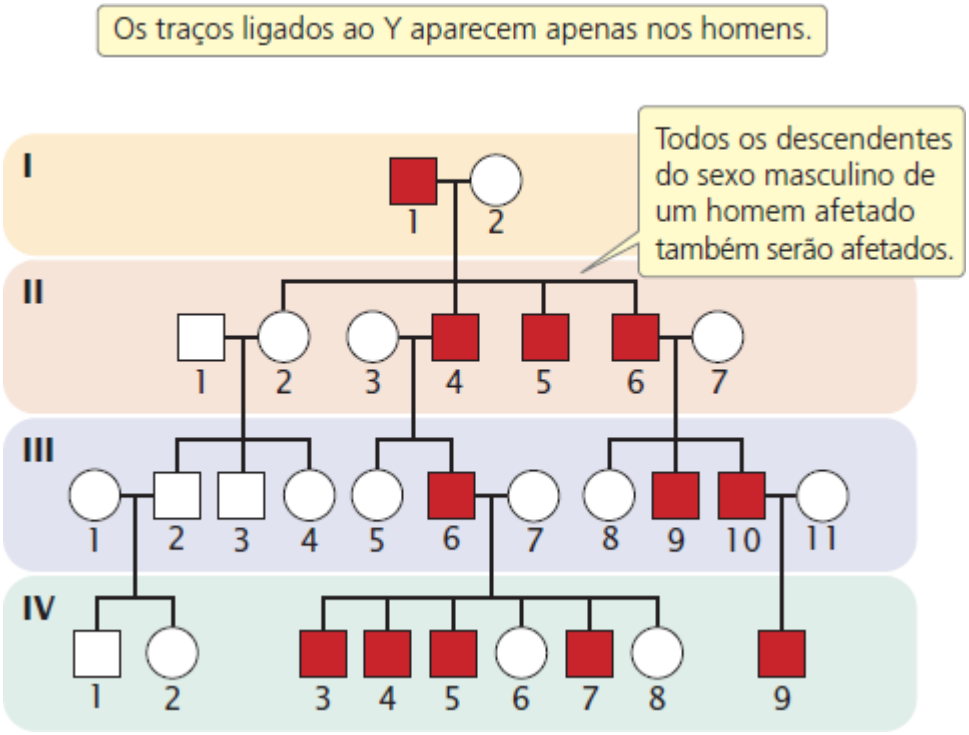


Figura 6.10 **Os traços ligados ao Y são encontrados apenas nos homens e são transmitidos do pai para todos os seus filhos.**

Quadro 6.1 Características no heredograma dos traços autossômicos recessivos, autossômicos dominantes, recessivos ligados ao X, dominantes ligados ao Y e ligados ao Y.

Traço autossômico recessivo

1. Em geral surge em ambos os sexos com a mesma frequência.
2. Tende a pular gerações.
3. Os indivíduos afetados em geral têm pais não afetados.
4. Quando os pais são heterozigotos, aproximadamente 25% dos filhos serão afetados.
5. Aparece com maior frequência em crianças de casamentos consanguíneos.

Traço autossômico dominante

1. Em geral surge em ambos os sexos com a mesma frequência.

2. Ambos os sexos transmitem o traço para seus descendentes.
3. Não pula gerações.
4. Os descendentes afetados têm necessariamente um genitor afetado, exceto se eles tiverem uma nova mutação.
5. Quando um dos genitores (heterozigoto) é afetado e o outro não, aproximadamente metade dos descendentes será afetada.
6. Os pais que não são afetados não transmitem o traço.

Traço recessivo ligado ao X

1. Em geral os homens são mais afetados que as mulheres.
2. Os homens afetados têm mães não afetadas, ou seja, o traço pula gerações.
3. Aproximadamente metade dos filhos de uma portadora (heterozigota) é afetada.
4. Nunca é transmitido do pai para o filho.
5. Todas as filhas de homens afetados são portadoras.

Traço dominante ligado ao X

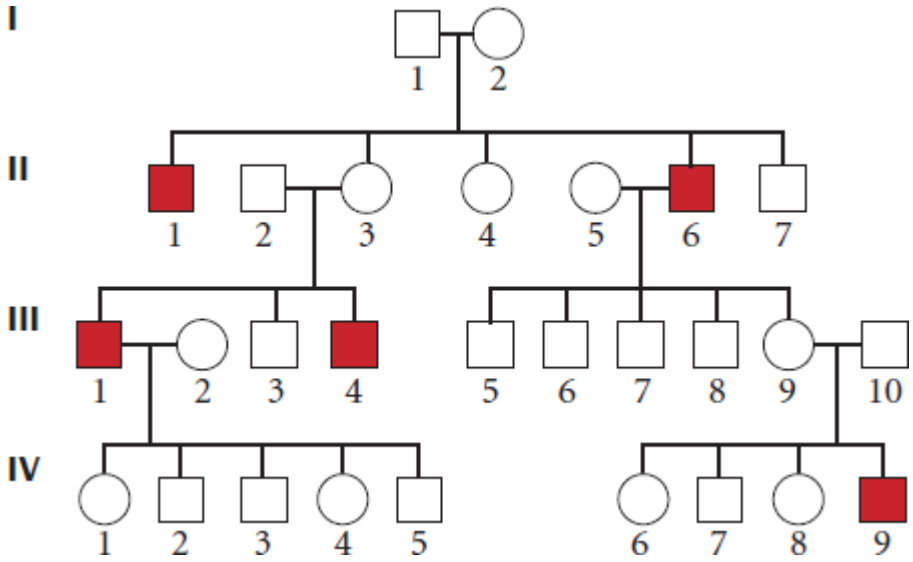
1. Homens e mulheres são afetados, mas, em geral, as mulheres são mais afetadas.
2. Não pula gerações. Os homens têm necessariamente a mãe afetada; as filhas afetadas devem ter a mãe ou o pai afetado.
3. Os homens afetados transmitirão o traço para todas as suas filhas.
4. As mães afetadas (se forem heterozigotas) transmitem o traço para metade de seus filhos e metade de suas filhas.

Traço ligado ao Y

1. Apenas os homens são afetados.
2. Transmitido do pai para todos os filhos.
3. Não pula gerações.

Problema desenvolvido

O heredograma a seguir representa a herança de um raro distúrbio em uma família estendida. Qual é a forma mais provável de herança para essa doença? (Considere que o traço exibe penetrância plena.)



Estratégias para a solução

De quais informações você precisa para solucionar o problema?

A forma mais provável de herança para o traço apresentado no heredograma.

Quais informações são fornecidas para solucionar o problema?

- O heredograma, que inclui informações sobre o sexo e as relações familiares das pessoas afetadas
- O traço é raro.

Etapas para a solução

Para responder a essa pergunta devemos considerar cada forma de herança e determinar qual, se alguma, podemos eliminar. O traço aparece apenas nos homens e as formas autossômicas dominantes e recessivas de herança são improváveis porque traços como esses aparecem igualmente em homens e mulheres. Além disso, a dominância autossômica pode ser eliminada porque algumas pessoas afetadas não têm um genitor afetado.

O traço é observado apenas nos homens nesse heredograma, o que poderia sugerir uma herança ligada ao Y. Entretanto, para um traço ligado ao Y, os homens afetados devem transmitir o traço para todos seus filhos, o que não é o caso aqui; II-6 é um homem afetado que tem quatro descendentes masculinos não afetados. Podemos eliminar a herança ligada ao Y.

A dominância ligada ao X pode ser eliminada porque os homens afetados devem transmitir um traço dominante ligado ao X para todas as suas filhas, e II-6 tem uma filha não afetada (III-9).

É mais comum os traços recessivos ligados ao X aparecerem nos homens, e os homens afetados têm mães portadoras não afetadas; o heredograma mostra esse padrão de herança. Para um traço ligado ao X, cerca de metade dos filhos de uma portadora heterozigota deve ser afetada. II-3 e III-9 são portadores suspeitos e cerca de metade de seus filhos (três de cinco) é afetada. Outra característica importante de um traço recessivo ligado ao X é que ele não é transmitido do pai para o filho. Observamos que não ocorre transmissão de pai para filho nesse heredograma. A forma recessiva ligada ao X é a forma mais provável de herança.

Para intensificar a prática, determine o modo de herança para os heredogramas no **Problema 24** ao final do capítulo.

6.3 O estudo sobre gêmeos e adoções pode ajudar a avaliar a importância dos genes e do meio ambiente

Os gêmeos e as adoções fornecem experimentos naturais para separar os efeitos dos genes e fatores ambientais ao determinar as diferenças nos traços. Essas duas técnicas são amplamente usadas em estudos genéticos.

Tipos de gêmeos

Existem dois tipos de gêmeos: **gêmeos dizigóticos** (não idênticos) surgem quando óvulos separados são fertilizados por dois

espermatozoides diferentes, produzindo zigotos geneticamente diferentes; os **gêmeos monozigóticos** (idênticos) surgem quando um único óvulo, fertilizado por um espermatozoide, divide-se em dois embriões separados no início do desenvolvimento.

Como os gêmeos monozigóticos surgem a partir de um único óvulo e espermatozoide, eles são idênticos do ponto de vista genético (exceto por algumas mutações somáticas raras) com 100% de seus genes em comum. Os gêmeos dizigóticos, por outro lado, têm em média apenas 50% de seus genes em comum, que é a mesma porcentagem que qualquer par de irmãos em comum tem. Como outros irmãos, os gêmeos dizigóticos podem ter o mesmo sexo ou não. A única diferença entre eles e outros irmãos é que os gêmeos têm a mesma idade e compartilharam o mesmo ambiente uterino. A tendência para nascimento de gêmeos dizigóticos tem, com frequência, caráter familiar e é influenciada pela hereditariedade e por fatores ambientais. Aparentemente, existe pouca tendência genética para produzir gêmeos monozigóticos.

Conceitos

Os gêmeos dizigóticos surgem a partir de dois óvulos fertilizados por dois espermatozoides distintos; em média, eles têm 50% de seus genes em comum. Os gêmeos monozigóticos surgem a partir de um único óvulo, fertilizado por um único espermatozoide, que se divide em dois embriões; eles têm 100% dos seus genes em comum.

Checagem dos conceitos 6

Por que os gêmeos monozigóticos são geneticamente idênticos, enquanto os gêmeos dizigóticos têm apenas 50% dos seus genes em comum?

a. Os gêmeos monozigóticos tendem a ser mais parecidos.
b. Os gêmeos monozigóticos se desenvolvem a partir de dois óvulos diferentes fertilizados pelo mesmo espermatozoide, enquanto os gêmeos dizigóticos se desenvolvem a partir de dois óvulos fertilizados por dois espermatozoides diferentes.
c. Os gêmeos monozigóticos se desenvolvem a partir de um único óvulo fertilizado por um espermatozoide, enquanto os gêmeos dizigóticos se desenvolvem a partir de um único óvulo fertilizado por dois espermatozoides diferentes.
d. Os gêmeos monozigóticos se desenvolvem a partir de um único óvulo fertilizado por um único espermatozoide, enquanto os gêmeos dizigóticos se desenvolvem a partir de dois óvulos fertilizados por dois espermatozoides diferentes.

Concordância em gêmeos

As comparações entre gêmeos dizigóticos e monozigóticos podem ser usadas para avaliar a importância da genética e dos fatores ambientais na produção de diferenças em uma característica. Essa avaliação é feita ao calcular a concordância para um traço. Se ambos os membros de um par de gêmeos têm um traço, é dito que eles são *concordantes*, se apenas um dos gêmeos tem o traço, eles são *discordantes*. A **concordância** é a porcentagem de pares de gêmeos que são concordantes para um traço. Como os gêmeos idênticos têm 100% dos seus genes em comum e os gêmeos dizigóticos têm em média apenas 50% em comum, os traços influenciados pela genética devem exibir maior concordância nos gêmeos monozigóticos. Por exemplo, quando um dos gêmeos monozigóticos tem epilepsia (**Quadro 6.2**), o outro tem epilepsia em 59% dos casos, então a concordância monozigótica para epilepsia é de 59%. Entretanto, quando um gêmeo dizigótico tem epilepsia, o outro tem epilepsia em apenas 19% das vezes (concordância dizigótica de 19%). A maior concordância nos gêmeos monozigóticos sugere que os genes influenciam a epilepsia, um achado com apoio dos resultados de estudos sobre a doença em outras famílias. Por outro lado, as taxas de concordância de morte por infecção aguda são semelhantes nos gêmeos monozigóticos e dizigóticos, o que sugere que a maioria das mortes por infecção tem pouca tendência hereditária. Os valores de concordância para vários traços humanos e doenças estão listados no Quadro 6.2.

Quadro 6.2 Concordância dos gêmeos monozigóticos e dizigóticos para vários traços.

Traço	Concordância (%)	
	Monozigótico	Dizigótico
1. Infarto do miocárdio (homens)	39	26
2. Infarto do miocárdio (mulheres)	44	14

3. Asma brônquica	47	24
4. Câncer (todos os órgãos)	12	15
5. Epilepsia	59	19
6. Morte por infecção aguda	7,9	8,8
7. Artrite reumatoide	32	6
8. Esclerose múltipla	28	5

Fontes: (*1 e 2*) B. Havald and M. Hauge, U.S. Public Health Service Publication 1103 (1963), pp. 61-67; (*3, 4, 5 e 6*) B. Havald and M. Hauge, Genetics and the Epidemiology of Chronic Diseases (U.S. Department of Health, Education, and Welfare, 1965); (*7*) J. S. Lawrence, Annals of Rheumatic Diseases 26:357-379, 1970; (*8*) G. C. Ebers et al., American Journal of Human Genetics 36:495, 1984.

A marca de uma influência genética em um traço específico é a maior concordância nos gêmeos monozigóticos comparados com os gêmeos dizigóticos. A alta concordância nos gêmeos monozigóticos por si só não é sinal de influência genética. Em geral, os gêmeos compartilham o mesmo ambiente – eles crescem no mesmo lar, têm os mesmos amigos, frequentam a mesma escola – e a alta concordância pode ser devida aos genes ou a um ambiente em comum. Se a alta concordância for causada por fatores ambientais, então, os gêmeos dizigóticos, que também compartilham o mesmo ambiente, devem ter uma concordância tão alta quanto a dos monozigóticos. Quando os genes influenciam o traço, entretanto, os pares de gêmeos monozigóticos devem exibir maior concordância que os dizigóticos, porque aqueles gêmeos têm maior porcentagem de genes em comum. É importante observar que qualquer discordância entre os gêmeos monozigóticos é causada por fatores ambientais, visto que eles são geneticamente idênticos. Por exemplo, no caso da epilepsia, a concordância dos gêmeos monozigóticos é consideravelmente menor que 100% (ver Quadro 6.2), o que sugere que, além das influências genéticas, os fatores ambientais também influenciam a variação no traço.

O uso de gêmeos na pesquisa genética se baseia na importante hipótese de que, quando a concordância para os gêmeos monozigóticos é maior que para os dizigóticos, é porque aqueles gêmeos são mais semelhantes em seus genes e não porque eles experimentaram um ambiente mais semelhante. Considera-se que o grau de semelhança ambiental entre os gêmeos monozigóticos e dizigóticos seja o mesmo. Essa suposição pode não estar sempre certa, sobretudo para o comportamento humano. Como eles são semelhantes, os gêmeos idênticos podem ser tratados de forma mais parecida pelos pais, professores e colegas que os gêmeos não idênticos. A evidência desse tratamento semelhante é observada na antiga tendência dos pais em vestir de forma parecida os gêmeos idênticos. Apesar dessa potencial complicação, os estudos com gêmeos têm um papel essencial no estudo da genética humana. **Resolva o Problema 30**

Estudo da asma em gêmeos

Para ilustrar o uso de gêmeos na pesquisa genética, vamos considerar um estudo sobre asma. A asma é caracterizada pela constrição das vias respiratórias e secreção de muco, provocando tosse e dificuldade respiratória. Os casos graves são potencialmente fatais. A asma é um importante problema de saúde nos países industrializados e parece estar aumentando. A incidência da doença na infância varia muito: algumas das taxas mais elevadas (de 21 a 27%) são encontradas na Austrália, no Reino Unido, na Suíça e no Brasil.

Sabemos que vários estímulos ambientais precipitam os ataques de asma, incluindo poeira, pólen, poluição do ar, infecções respiratórias, exercício, ar frio e estresse emocional. Em muitos casos, as alergias acompanham a asma, o que sugere que a asma seja um distúrbio do sistema imunológico, mas a relação exata entre a função imunológica e a asma é pouco compreendida. Vários estudos mostram que os fatores genéticos são importantes na asma.

Um estudo genético sobre asma na infância foi realizado como parte do Twins Early Development Study na Inglaterra, um projeto de pesquisa contínuo que estuda mais de 15.000 gêmeos nascidos no Reino Unido entre 1994 e 1996. Esses gêmeos foram submetidos à avaliação quanto a linguagem, desenvolvimento cognitivo, problemas comportamentais e desempenho na escola entre 7 e 9 anos e foram examinadas as contribuições genéticas e ambientais para vários traços. No estudo sobre asma, os pesquisadores examinaram uma amostra de 4.910 gêmeos com 4 anos. Foi perguntado aos pais dos gêmeos se havia prescrição de medicação para controle da asma; as crianças que receberam medicação para asma foram consideradas asmáticas.

O valor de concordância para os gêmeos monozigóticos (65% em 1.658 pares de gêmeos) foi muito maior que o valor para os gêmeos dizigóticos (37% em 3.252 pares de gêmeos), e os pesquisadores concluíram que, entre as crianças de 4 anos incluídas no estudo, a asma era fortemente influenciada pelos fatores genéticos. O fato de que até os gêmeos monozigóticos eram discordantes 35% das vezes indica que fatores ambientais também atuam no desenvolvimento da asma.

Conceitos

A maior concordância para os gêmeos monozigóticos comparada com a concordância para gêmeos dizigóticos indica que os fatores genéticos influenciam as diferenças em um traço. Uma concordância menor que 100% para gêmeos monozigóticos indica que os fatores ambientais têm um papel significativo.

Checagem dos conceitos 7

Um traço exibe 100% de concordância para os gêmeos monozigóticos e dizigóticos. Que conclusão você pode tirar sobre o papel dos fatores genéticos na determinação das diferenças no traço?

a. Os fatores genéticos são muito importantes.
b. Os fatores genéticos são pouco importantes.
c. Os fatores genéticos não são importantes.
d. Os fatores genéticos e ambientais são importantes.

Estudos sobre adoção

Outra técnica importante usada pelos geneticistas para estudar a herança humana é o estudo das pessoas adotadas. Essa abordagem é uma das mais importantes para diferenciar os efeitos dos genes e ambientes nas características das pessoas.

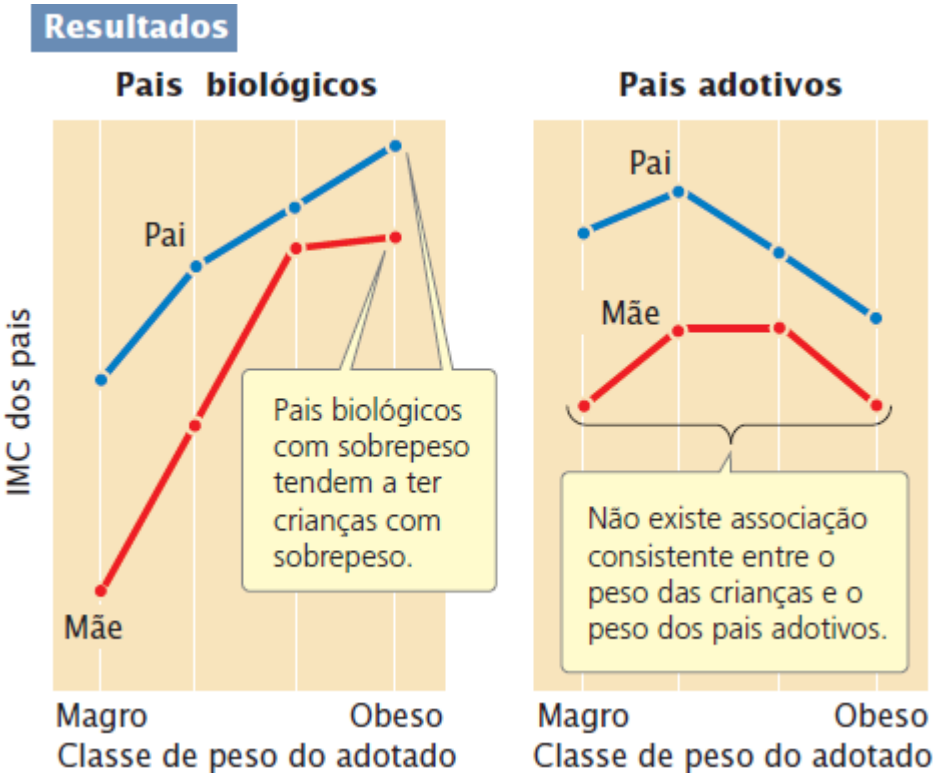
Por motivos diversos, muitas crianças são separadas de seus pais biológicos logo após o nascimento e adotadas por adultos com os quais elas não têm nenhuma relação genética. As pessoas adotadas não têm mais genes em comum com seus pais adotivos do que com duas pessoas escolhidas aleatoriamente, entretanto, elas compartilham o ambiente com seus pais adotivos. Em contrapatida, as pessoas adotadas têm 50% dos seus genes em comum com cada um dos seus pais biológicos, mas não compartilham o mesmo ambiente com eles. Se as pessoas adotadas e seus pais adotivos mostram semelhanças em uma característica, elas podem ser atribuídas aos fatores ambientais. Se, por outro lado, as pessoas adotadas e seus pais biológicos apresentam semelhanças, é provável que elas sejam causadas por fatores genéticos. As comparações das pessoas adotadas com seus pais adotivos e seus pais biológicos podem, portanto, ajudar a definir os papéis dos fatores genéticos e ambientais na determinação da variação humana. Por exemplo, os estudos de adoção foram úteis ao mostrar que a esquizofrenia tem uma base genética. Estudos sobre adoção também mostraram que a obesidade, medida como o índice de massa corporal, é pelo menos em parte, influenciada pela genética (**Figura 6.11**).

Experimento

Pergunta: O índice de massa corporal (IMC) é influenciado por fatores genéticos?

Métodos

Compare o índice de massa corporal das crianças adotadas com o dos pais adotivos e biológicos.



Conclusão: Fatores genéticos influenciam o índice de massa corporal.

Figura 6.11 Os estudos sobre adoção demonstram que a obesidade tem influência genética. (Redesenhada com permissão do New England Journal of Medicine 314:195, 1986.)

Essas pesquisas pressupõem que os ambientes das famílias biológicas e adotivas sejam independentes (ou seja, não mais do que seria esperado por acaso). Tal suposição pode não estar sempre correta, porque as agências de adoção escolhem com cuidado os pais adotivos e podem selecionar uma família que lembre a família biológica. Assim, algumas das semelhanças entre as pessoas adotadas e seus pais biológicos podem ser causadas por esses ambientes semelhantes e não por causa dos fatores genéticos comuns. Além disso, os descendentes e a mãe biológica compartilham o mesmo ambiente durante o desenvolvimento pré-natal. Resolva o Problema 33

Conceitos

As semelhanças entre as pessoas adotadas e seus pais adotivos sem relação genética indicam que os fatores ambientais afetam uma característica específica; as semelhanças entre as pessoas adotadas e seus pais biológicos indicam que os fatores genéticos influenciam a característica.

Checagem dos conceitos 8

Quais suposições sustentam o uso de estudos sobre adoção na genética?

a. Os adotados não têm contato com seus pais biológicos após o nascimento.
b. Os pais adotivos e os pais biológicos não são parentes.
c. Os ambientes dos pais biológicos e dos pais adotivos são independentes.
d. Todas as opções anteriores.

6.4 O aconselhamento e os exames genéticos fornecem informações às pessoas preocupadas com doenças e traços genéticos

Nosso conhecimento sobre doenças e distúrbios genéticos humanos vem se expandindo rapidamente nos últimos anos. O catálogo *Online Mendelian Inheritance in Man* lista atualmente mais de 21.000 distúrbios, doenças genéticas, genes e traços humanos que têm uma base genética simples. A pesquisa forneceu numerosas informações sobre herança, localização no cromossomo, base bioquímica e sintomas de muitos traços, doenças e distúrbios genéticos. Essas informações são úteis para ajudar as pessoas que têm uma condição genética.

Aconselhamento genético

O **aconselhamento genético** é um campo que fornece informações a pacientes e outras pessoas que se preocupam com questões relacionadas à hereditariedade. É um processo educacional que ajuda pacientes e membros da família a lidar com muitos aspectos de uma condição genética, incluindo diagnóstico, informações sobre sintomas e tratamento e informações sobre o modo de herança. O aconselhamento genético também ajuda essas pessoas a lidar com o estresse psicológico e físico que está associado ao distúrbio. Evidentemente, todas essas questões não podem ser discutidas por uma única pessoa, então grande parte do aconselhamento genético é feito por uma equipe que pode incluir psicólogos, médicos, geneticistas e pessoal de laboratório. O **Quadro 6.3** lista alguns motivos comuns para procurar o aconselhamento genético.

Quadro 6.3 Motivos comuns para procurar o aconselhamento genético.

1. A pessoa sabe de uma doença genética na família.
2. Um casal tem uma criança com uma doença genética, defeito congênito ou anomalia cromossômica.
3. Um casal tem uma criança com déficit intelectual ou um parente próximo intelectualmente incapaz.
4. Uma mulher mais velha engravida ou quer engravidar. Não existe um acordo sobre a idade na qual uma futura mãe, que não tem outro fator de risco, deve procurar aconselhamento genético, mas muitos especialistas sugerem 35 anos ou mais.
5. Marido e esposa são parentes (p. ex., primos de primeiro grau).
6. Um casal está com dificuldade para engravidar.
7. Uma grávida está preocupada sobre a exposição a uma substância no ambiente (fármaco, substância química, vírus) que provoca defeitos congênitos.
8. Um casal precisa de ajuda para interpretar os resultados de um teste pré-natal ou outro teste.
9. Ambos os futuros pais são portadores conhecidos de uma doença genética recessiva ou ambos pertencem a um grupo étnico com alta frequência de uma doença genética.

Em geral, o aconselhamento genético inicia com o diagnóstico da condição. Com base no exame físico, testes bioquímicos, exame de DNA, análise cromossômica, história familiar e outras informações, o médico determina a causa. Um diagnóstico exato é crítico, porque o tratamento e a probabilidade de transmitir a condição podem variar, dependendo do diagnóstico. Por exemplo, existem vários tipos diferentes de nanismo, que podem ser causados por anomalias cromossômicas, mutações de um único gene, desequilíbrios hormonais ou fatores ambientais. Pessoas com nanismo resultante de um gene autossômico dominante têm chance de 50% de transmitir essa característica para seus filhos, enquanto as pessoas cujo nanismo é causado por um raro gene recessivo têm baixa probabilidade de transmitir o traço.

Quando a natureza da condição é conhecida, um conselheiro genético faz uma reunião com o paciente e membros da família e explica o diagnóstico. Um heredograma da família pode ser construído e a probabilidade de transmitir a condição para gerações futuras pode ser calculada para diferentes membros da família. O conselheiro auxilia a família a interpretar os riscos genéticos e explica as várias opções de reprodução disponíveis, incluindo diagnóstico pré-natal, inseminação artificial e fertilização *in vitro*. Muitas vezes, os membros da família têm dúvidas sobre o exame genético que pode estar disponível para determinar se eles carreiam uma mutação. O conselheiro os ajuda a decidir se o exame genético é adequado e quais testes buscar. Após receber os resultados dos testes, o conselheiro ajuda a interpretá-los.

A decisão de uma família sobre futuras gestações depende da magnitude do risco genético, da gravidade e dos efeitos da condição, da importância de ter filhos e das questões religiosas e culturais. Tradicionalmente, os conselheiros genéticos são treinados para aplicar o aconselhamento *não direcionado*, o que significa que eles fornecem as informações e facilitam a discussão, mas não emitem suas opiniões e valores para a discussão. O objetivo do aconselhamento não direcionado é que a família alcance sua própria decisão com base nas melhores informações disponíveis.

Devido ao número crescente de testes genéticos e a complexidade da avaliação do risco genético, existe atualmente um movimento no sentido contrário do aconselhamento não direcionado. O objetivo ainda é fornecer informações à família sobre todas as opções e alcançar a melhor decisão, mas pode ser necessário que o conselheiro, em alguns casos, recomende algumas opções, como um médico recomenda os tratamentos mais adequados para seu paciente.

Quem faz o aconselhamento genético? Nos EUA, cerca de 6.000 profissionais de saúde estão atualmente certificados em genética pelo American Board of Medical Genetics ou o American Board of Genetic Counseling. Cerca de metade deles é treinada especificamente no aconselhamento genético, em geral ao terminar um programa de mestrado de 2 anos que fornece educação sobre genética e aconselhamento. A maior parte dos outros profissionais é de médicos e cientistas certificados em genética médica e clínica. Por causa da carência de conselheiros em genética e geneticistas, as informações sobre exame e risco genético são fornecidas por médicos de atenção básica, enfermeiras e assistentes sociais. Resolva o Problema 10

Conceitos

O aconselhamento genético é um processo educacional que fornece informações para os pacientes e suas famílias sobre uma condição genética, suas implicações clínicas, a forma de herança e as opções de reprodução.

Exame genético

O principal objetivo do exame genético é identificar o potencial para uma condição genética em um estágio inicial. Em alguns casos, o exame genético permite que as pessoas façam escolhas informadas sobre a reprodução. Em outros casos, ele permite a intervenção precoce que pode diminuir ou até evitar o desenvolvimento da condição. Para quem sabe que corre risco de desenvolver uma condição genética, esse exame pode ajudar a aliviar a ansiedade associada à incerteza da sua situação. O exame genético inclui exames pré e pós-natal.

Os testes genéticos pré-natal são realizados antes do nascimento e atualmente incluem procedimentos para diagnóstico de várias doenças e distúrbios genéticos (**Quadro 6.4**). O principal objetivo desses testes é fornecer informações de que a família pode precisar para tomar decisões durante a gestação e, em alguns casos, preparar-se para o nascimento de uma criança com uma condição genética. Várias abordagens para o diagnóstico pré-natal estão descritas nas seções a seguir.

Ultrassonografia. Algumas condições genéticas podem ser detectadas pela visualização direta do feto. É comum que isso seja feito pela **ultrassonografia**, também chamada de ultrassom. Nessa técnica, um som de alta frequência é direcionado para o útero; quando as ondas sonoras encontram tecido denso, elas voltam e são transformadas em uma imagem. Pode ser determinado o tamanho do feto, assim como as condições genéticas como defeitos do tubo neural (defeitos no desenvolvimento da medula espinal e crânio) e anomalias no esqueleto. A ultrassonografia é um procedimento padrão feito durante a gestação para estimar a idade do feto, determinar seu sexo e verificar se há distúrbios de desenvolvimento ou outros problemas.

Quadro 6.4 Exemplos de doenças e distúrbios genéticos que podem ser detectados no pré-natal e as técnicas usadas em sua detecção.

Distúrbio	Método de detecção
Anomalias cromossômicas	Exame de um cariótipo a partir das células obtidas por amniocentese ou amostragem das vilosidades coriônicas
Fendas palatina e labial	US
Fibrose cística	Análise de DNA das células obtidas por amniocentese ou amostragem das vilosidades coriônicas
Nanismo	US ou radiografia; algumas formas podem ser detectadas por análise de DNA das células obtidas por amniocentese ou amostragem das vilosidades coriônicas
Hemofilia	Amostragem do sangue fetal* ou análise de DNA das células obtidas por amniocentese ou amostragem das vilosidades coriônicas
Síndrome de Lesch-Nyhan	Testes bioquímicos com as células obtidas por amniocentese ou amostragem das vilosidades coriônicas
Defeitos do tubo neural	Triagem inicial com teste do sangue materno, seguido por testes bioquímicos no líquido amniótico obtido por amniocentese ou por detecção de defeitos

	congênitos por US
Osteogênese imperfeita	US ou radiografia
Fenilcetonúria	Análise de DNA das células obtidas por amniocentese ou amostragem das vilosidades coriônicas
Anemia falciforme	Amostragem do sangue fetal* ou análise de DNA das células obtidas por amniocentese ou amostragem das vilosidades coriônicas
Doença de Tay-Sachs	Testes bioquímicos com as células obtidas por amniocentese ou amostragem das vilosidades coriônicas

*Uma amostra do sangue fetal é obtida ao inserir uma agulha no cordão umbilical. *US*, ultrassonografia.

Amniocentese. O tradicional teste pré-natal requer tecido fetal, que pode ser obtido de várias formas. O método mais usado é a **amniocentese**, um processo para obter uma amostra de líquido amniótico da gestante (**Figura 6.12**). O líquido amniótico – a substância que preenche o saco amniótico e circunda o feto em desenvolvimento – tem células do feto que podem ser usadas para exame genético.

A amniocentese é feita na rotina como um procedimento ambulatorial com ou sem o uso de anestésico local (ver Figura 6.12). Os testes genéticos são, então, realizados nas células cultivadas. As complicações com a amniocentese (em sua grande parte o aborto) são raras, apenas 1 em 400 procedimentos.

Amostragem das vilosidades coriônicas. A principal desvantagem da amniocentese é que ela é realizada entre a 15ª e a 18ª semana de gestação (embora alguns obstetras consigam fazer a amniocentese antes desse período). As células obtidas por amniocentese têm de ser cultivadas antes de os testes genéticos serem feitos, o que requer mais tempo. Por esses motivos, as informações genéticas do feto não estão disponíveis antes da 17ª ou 18ª semana de gestação. Um aborto nesse estágio tem o risco de complicações e é mais estressante para os pais. A **amostragem das vilosidades coriônicas** pode ser feita mais cedo (entre a 10ª e a 12ª semana de gestação) e coleta mais tecido fetal, o que elimina a necessidade de cultivo das células.

Nesse exame, um cateter – um tubo plástico mole – é colocado em contato com o cório, a camada externa da placenta (**Figura 6.13**). É aplicada sucção e um pequeno pedaço de cório é removido. Embora o cório seja composto de células fetais, ele é uma parte da placenta que é expelida do útero após o parto. O tecido removido não é de fato oriundo do feto, mas tem milhões de células em divisão que podem ser usadas diretamente em muitos testes genéticos. Há maior risco de complicação que a amniocentese, e os resultados de vários estudos sugerem que esse procedimento pode aumentar a incidência de defeitos nos membros do feto quando feito antes da 10ª semana de gestação.

As células fetais obtidas por amniocentese ou amostragem de vilosidades coriônicas podem ser usadas para preparar um **cariótipo**, que é uma imagem de um conjunto completo dos cromossomos na metáfase. Os cariótipos podem ser estudados para anomalias cromossômicas (ver *Capítulo 8*). As análises bioquímicas podem ser feitas nas células fetais para determinar a existência de produtos metabólicos particulares dos genes. Para doenças genéticas nas quais a sequência de DNA do gene responsável foi determinada, a sequência de DNA (exame de DNA, ver *Capítulo 19*) pode ser examinada quanto aos alelos defeituosos.

Testes de triagem no sangue materno. O maior risco de algumas condições genéticas pode ser detectado ao examinar os níveis de algumas substâncias no sangue da mãe (chamado de **teste de triagem no sangue materno**). Entretanto, esses testes não determinam se há algum problema genético, mas sim indicam que o feto corre maior risco e, por isso, são chamados de testes de *triagem*. Quando é detectado um risco maior, em geral são realizados testes de acompanhamento (testes de triagem sanguínea adicionais, US, amniocentese ou os três).

Uma substância examinada nos testes de triagem materna é a α-fetoproteína, uma proteína que é normalmente produzida pelo feto durante o desenvolvimento e está presente no sangue fetal, no líquido amniótico e no sangue da mãe durante a gravidez. O nível da α-fetoproteína é muito maior que o normal quando o feto tem um defeito do tubo neural ou um entre vários outros distúrbios. Algumas anomalias cromossômicas produzem níveis de α-fetoproteína abaixo do normal. Medir a quantidade de α-fetoproteína no sangue da mãe fornece uma indicação dessas condições.

O American College of Obstetricians and Gynecologists recomenda que os médicos ofereçam os testes de triagem no sangue materno a todas as gestantes. Um teste típico, realizado entre 11 e 13 semanas de gestação, determina o nível de gonadotrofina coriônica humana (hCG, um hormônio da gravidez) e uma substância chamada de proteína plasmática A, associada à gravidez (PAPP-A). Quando o feto tem alguns defeitos cromossômicos, o nível de PAPP-A tende a ser menor e o nível de hCG, maior. O

risco de uma anomalia cromossômica é calculado com base nos níveis de hCG e PAPP-A no sangue da mãe, junto com os resultados da ultrassonografia. Outro teste mede os níveis de quatro substâncias: α-fetoproteína, hCG, estriol e inibina. O risco de anomalias cromossômicas e alguns tipos de defeitos congênitos é calculado com base nos níveis combinados das quatro substâncias mais a idade, peso, histórico étnico e número de fetos da mãe. Essa abordagem detecta com sucesso a síndrome de Down (graças às três cópias do cromossomo 21) em 81% das vezes.

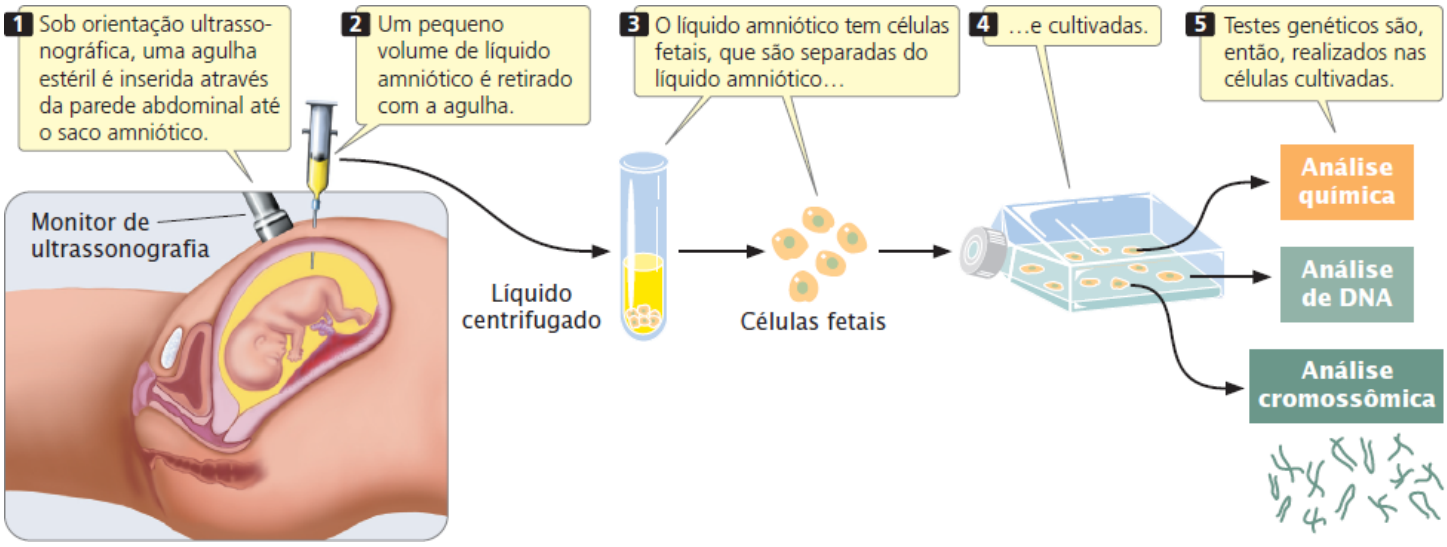


Figura 6.12 A amniocentese é um procedimento usado para obter células fetais para exame genético.

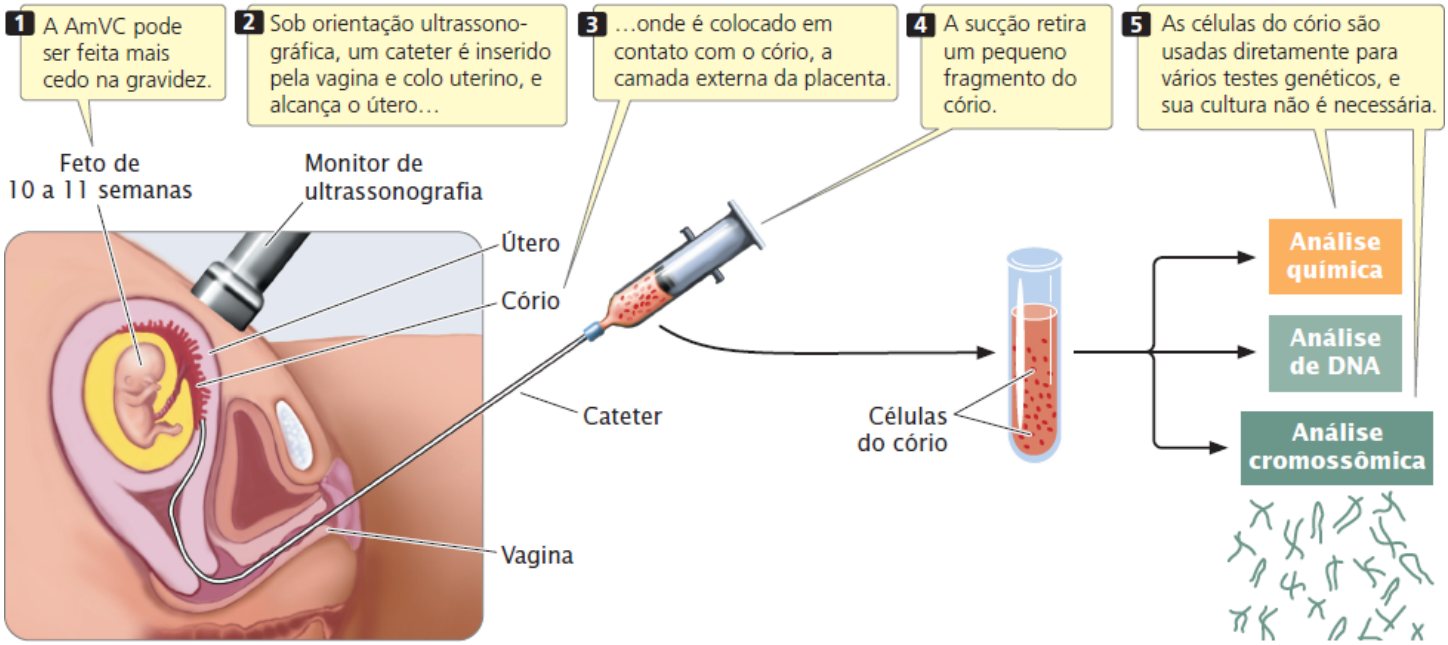


Figura 6.13 A amostragem das vilosidades coriônicas (AmVC) é outro procedimento usado para obter células fetais para exame genético.

Diagnóstico genético pré-natal não invasivo. Os testes de pré-natal que usam apenas o sangue materno são interessantes porque são não invasivos e não apresentam risco para o feto. Além dos testes de triagem no sangue materno, que medem as substâncias químicas produzidas pelo feto ou pela placenta, os procedimentos chamados de **diagnóstico genético pré-natal não invasivo** examinam diretamente o DNA fetal no sangue da mãe. Esses testes podem ser realizados antes de 10 semanas após a concepção.

Durante a gravidez, algumas células fetais são liberadas para o sistema circulatório da mãe, onde se misturam e circulam com seu sangue. Avanços recentes permitem detectar e separar as células fetais das células do sangue materno (um processo chamado de **seleção de célula fetal**) com o uso de *laser* e máquinas automatizadas para separação de células. As células fetais obtidas podem ser cultivadas para análise cromossômica ou usadas como uma fonte de DNA fetal para exame molecular (ver *Capítulo 19*). O sangue materno também tem fragmentos livremente flutuantes de DNA fetal, que são liberados quando as células fetais são degradadas. O DNA fetal pode ser sequenciado e testado para verificar mutações. Também estão disponíveis testes para determinar o número de cópias das variantes genéticas, permitindo a determinação do número de cromossomos

carreados pelo feto, de modo que as anomalias cromossômicas como a síndrome de Down possam ser detectadas a partir do DNA fetal. O diagnóstico genético pré-natal não invasivo está sendo usado atualmente para determinar o tipo de sangue do feto, detectar a síndrome de Down e outros distúrbios cromossômicos e para identificar mutações para doenças genéticas como fibrose cística e talassemia (um distúrbio sanguíneo). Essa tecnologia cria o potencial para o uso de uma única amostra de sangue da mãe para testar várias doenças genéticas e até traços comuns no feto. Essa possibilidade aumenta o número de questões sociais e éticas sobre o uso de tal informação nas decisões sobre reprodução.

Diagnóstico genético pré-implante. Os testes genéticos pré-natal fornecem aos futuros pais muitas informações sobre a saúde de seus futuros filhos. Novas tecnologias reprodutivas fornecem opções aos casais para usar essas informações. Uma dessas tecnologias é a fertilização *in vitro*, procedimento no qual hormônios são usados para induzir a ovulação. Os óvulos maduros são retirados cirurgicamente da superfície do ovário, colocados em uma placa e fertilizados com espermatozoides. O embrião resultante é então implantado no útero. Já nasceram milhares de bebês oriundos da fertilização *in vitro.*

O exame genético pode ser combinado com esse tipo de fertilização para permitir o implante dos embriões que estejam livres de um defeito genético específico. Chamado de **diagnóstico genético pré-implante** (PGD), essa técnica permite que as pessoas portadoras de um defeito genético evitem ter uma criança com a doença. Por exemplo, se uma mulher é portadora de uma doença recessiva ligada ao X, é esperado que aproximadamente metade de seus filhos tenha a doença. Por meio da fertilização *in vitro* e exame pré-implante, um embrião sem o distúrbio pode ser selecionado para implante em seu útero.

O procedimento começa com a produção de vários embriões de uma célula por meio da fertilização *in vitro*. Os embriões se dividem várias vezes até alcançarem o estágio de 8 ou 16 células. Neste momento, uma célula é removida de cada embrião e é feito o teste para detectar anomalia genética. A remoção de uma única célula nesse estágio inicial não é prejudicial ao embrião. Após determinar quais embriões não têm o distúrbio, um embrião saudável é selecionado e implantado no útero da mulher.

O diagnóstico genético pré-implante requer a capacidade de realizar um teste genético em uma única célula. Esse exame é possível graças ao uso da reação em cadeia da polimerase, pela qual pequenas quantidades de DNA podem ser amplificadas (replicadas) rapidamente (ver *Capítulo 19*). Após a amplificação do DNA da célula, a sequência de DNA é examinada. O diagnóstico genético pré-implante resultou no nascimento de milhares de crianças saudáveis. Seu uso levanta várias questões éticas, porque ele pode ser usado para selecionar ou descartar traços genéticos que não têm relação com questões clínicas. Por exemplo, ele pode ser usado para selecionar uma criança com genes para uma cor de olhos ou genes para estatura maior.

Triagem do recém-nascido. O exame para distúrbios genéticos em recém-nascidos é chamado de **triagem do recém-nascido** (teste do pezinho). Todos os estados nos EUA e muitos outros países exigem por lei que sejam realizados testes para algumas doenças e condições genéticas em recém-nascidos. Em 2006, o American College of Medical Genetics recomendou a triagem obrigatória para 29 condições (**Quadro 6.5**) e muitos estados adotaram essa lista para o exame em recém-nascidos. Essas condições genéticas foram escolhidas porque a identificação precoce pode levar ao tratamento eficiente. Por exemplo, como foi mencionado no Capítulo 5, a fenilcetonúria é uma doença autossômica recessiva que, se não for tratada cedo, pode levar à incapacidade intelectual. Mas a intervenção precoce, com uma dieta modificada, evita a evolução da doença.

Quadro 6.5 Condições genéticas recomendadas para a triagem obrigatória segundo American College of Medical Genetics.

Deficiência de acil-coA desidrogenase de cadeia média
Hipotireoidismo congênito
Fenilcetonúria
Deficiência de biotinidase
Anemia falciforme (doença da Hb SS)
Hiperplasia suprarrenal congênita (deficiência da 21-hidroxilase)
Acidemia isovalérica
Deficiência de acil-coA desidrogenase de cadeia muito longa
Doença (da urina) em xarope de bordo
Galactosemia clássica

β-talassemia Hb S
Doença Hb S C
Deficiência da L-3-hidroxiacil-CoA desidrogenase de cadeia longa
Acidemia glutárica do tipo I
Acidúria 3-hidroxi-3-metilglutárica
Deficiência de proteína trifuncional
Deficiência múltipla da carboxilase
Acidemia metilmalônica (deficiência de mutase)
Homocistinúria (devida a deficiência de cistationina β sintase)
Deficiência de 3-metilcrotonil-CoA carboxilase
Perda auditiva
Acidemia metilmalônica
Acidemia propiônica
Defeito na captação da carnitina
Deficiência de β-cetotiolase
Citrulinemia
Acidemia argininossuccínica
Tirosinemia do tipo I
Fibrose cística

Exame pré-sintomático. Além da investigação de doenças genéticas nos fetos e recém-nascidos, atualmente é possível fazer exame em adultos saudáveis para pesquisar genes que podem predispô-los a uma condição genética. Esse tipo de exame é conhecido como **exame genético pré-sintomático**. Por exemplo, ele está disponível para membros de famílias que têm uma forma autossômica dominante do câncer de mama. Nesse caso, a identificação precoce do alelo responsável pela doença permite a observação e a detecção precoce dos tumores. O exame pré-sintomático também está disponível para algumas doenças genéticas para as quais não existe tratamento, como a doença de Huntington, uma doença autossômica dominante que resulta em lenta deterioração física e mental em pessoas de meia-idade.

Triagem de heterozigoto. Outra forma de exame genético nos adultos é a **triagem de heterozigoto**. Nesse tipo de triagem, os membros de população são testados para identificar os portadores heterozigotos dos alelos responsáveis por doença recessiva, ou seja, pessoas que são saudáveis, mas podem ter filhos com uma doença específica.

O exame para a doença de Tay-Sachs é um exemplo de sucesso de triagem de heterozigoto. Na população da América do Norte, a frequência dessa doença é de apenas 1 pessoa em 360.000. Entre os judeus Asquenaze (descendentes da população judia oriunda da Europa oriental e central), a frequência é 100 vezes maior. Um simples teste de sangue é usado para identificar os judeus Asquenaze portadores do alelo para a doença de Tay-Sachs. Se um homem e uma mulher forem heterozigotos, prevê-se que aproximadamente um em quatro de seus filhos terá a doença. Os casais identificados como portadores heterozigotos podem usar a informação para decidir se terão filhos. Um teste pré-natal também está disponível para determinar se o feto de um casal de risco terá a doença. Os programas de triagem levam a um importante declínio no número de crianças com ascendência Asquenaze com a doença de Tay-Sachs (atualmente menos de 10 crianças por ano nos EUA). **Resolva o Problema 12**

Conceitos

O exame genético é usado para rastrear doenças genéticas em recém-nascidos, detectar pessoas que são

heterozigotas para doenças recessivas, detectar alelos responsáveis por doenças em quem ainda não desenvolveu os sintomas e detectar alelos defeituosos em bebês não nascidos. O diagnóstico genético pré-implante, combinado com a fertilização *in vitro*, possibilita a seleção de embriões que estão livres de doenças genéticas específicas.

Checagem dos conceitos 9

Qual é a diferença entre o diagnóstico genético pré-implante e o exame genético pré-natal?

Como interpretar os testes genéticos

Atualmente, existem milhares de testes genéticos disponíveis e muitos outros estarão disponíveis após pesquisas clínicas. A pesquisa futura aumentará muito o número e a complexidade dos testes disponíveis. Muitos deles serão indicados para complexas doenças multifatoriais que são influenciadas pela genética e pelo ambiente, como coronariopatia, diabetes, asma, alguns tipos de câncer e depressão.

A interpretação dos resultados dos testes genéticos é complicada por vários fatores. Em primeiro lugar, algumas doenças genéticas são causadas por várias mutações diferentes. Por exemplo, mais de 1.000 diferentes mutações em um único *locus* podem provocar fibrose cística, uma doença autossômica recessiva na qual o transporte de íon cloreto é defeituoso (ver *Capítulo 5*). Tipicamente, testes genéticos pesquisam apenas as mutações mais comuns; as mutações incomuns e raras não são detectadas. Portanto, um resultado negativo não significa que um defeito genético esteja ausente; indica apenas que a pessoa não tem uma mutação comum. Quando membros da família têm a doença, seu DNA pode ser examinado para determinar a natureza da mutação, e outros membros podem ser, então, testados para a mesma mutação, mas essa opção não é possível se os membros afetados não estão disponíveis ou relutam a serem testados.

Um segundo problema está na interpretação dos resultados. Para uma doença genética clássica, como a doença de Tay-Sachs, a herança das duas cópias do gene praticamente garante que a pessoa terá a doença. Entretanto, esse não é o caso para muitas doenças genéticas nas quais a penetrância é incompleta e os fatores ambientais têm um papel importante. Para essas condições, ter uma mutação que predispõe a uma doença apenas eleva o risco de a pessoa adquiri-la. O risco associado a uma mutação específica é uma estimativa estatística, com base no efeito geral da mutação em muitas pessoas. Nesse caso, o risco calculado pode fornecer pouca informação útil para uma pessoa. Também é importante ter em mente que não existem testes para muitos traços e distúrbios genéticos.

Conceitos

Interpretar os testes genéticos é complicado por haver múltiplas mutações, penetrância incompleta e influência de fatores ambientais.

Testes genéticos de coleta domiciliares

Hoje em dia são oferecidos vários testes genéticos para qualquer pessoa interessada em investigar sua hereditariedade, sem precisar de um profissional de saúde. Esses **testes genéticos de coleta domiciliares** estão disponíveis para verificar um grande e crescente conjunto de condições genéticas em adultos e crianças, desde distúrbios de um único gene, como a fibrose cística, a condições multifatoriais como obesidade, doença cardiovascular, desempenho atlético e predisposição para dependência a nicotina. Esses testes também estão disponíveis para testar paternidade e determinar seus antepassados.

Muitos testes genéticos domiciliares são anunciados e vendidos pela internet. Uma pessoa pede um teste e a empresa envia um *kit* para coletar uma amostra de DNA (geralmente as células da saliva da cavidade oral ou uma amostra de sangue). A pessoa coleta a amostra e envia para a empresa, que faz o teste e envia o resultado de volta. Os geneticistas, agentes de saúde pública e defensores dos consumidores levantaram várias questões sobre esse tipo de teste, incluindo a preocupação de que alguns testes são oferecidos sem informações adequadas e aconselhamento genético e que os consumidores não estão preparados para interpretar os resultados. Outras questões se voltam para a exatidão de alguns testes, sua confidencialidade e se as indicações de risco fornecidas pelo teste são mesmo úteis.

Os defensores desses testes afirmam que eles possibilitam um maior acesso a exames e garantem a confidencialidade. Muitos estados nos EUA não regulam este tipo de teste e, atualmente, existe pouca fiscalização no âmbito federal. **Resolva o Problema 15**

Discriminação genética e privacidade

Com o desenvolvimento de muitos novos testes genéticos, surgiram questões sobre a privacidade das informações e o potencial para a discriminação genética. A pesquisa mostra que muitas pessoas com risco para doenças genéticas evitam o exame porque têm medo de que os resultados dificultem seu acesso ao seguro de saúde ou afetem de maneira adversa sua chance de conseguir um emprego. Algumas pessoas que procuram o exame genético pagam por contra própria e usam pseudônimos para evitar que os resultados sejam inseridos em seus prontuários. O medo da discriminação genética é reforçado por práticas antigas. Nos anos 1970, alguns afro-americanos identificados como portadores da anemia falciforme (um distúrbio autossômico recessivo) encararam a discriminação nos seus empregos e tiveram dificuldade para conseguir o seguro de saúde, apesar de serem portadores saudáveis.

Em resposta a essas questões, o Congresso norte-americano aprovou o **Genetic Information Nondiscrimination Act** (GINA) em 2008. Esse ato proíbe que as seguradoras de saúde usem informações genéticas para tomar decisões sobre coberturas e valores de seguro saúde. Ele também evita que empregadores usem informações genéticas nos exames admissionais e proíbe as seguradoras de saúde e empregadores que peçam ou exijam que uma pessoa seja submetida a um teste genético. Os resultados do exame genético receberam proteção por outros regulamentos federais que cobrem o uso e a revelação de informações de saúde da pessoa. Entretanto, o GINA cobre apenas o seguro-saúde e o emprego; ele não se aplica à vida, à incapacidade e à assistência a cuidados a longo prazo. **Resolva o Problema 16**

Conceitos

O número crescente de testes genéticos e sua complexidade levantaram várias questões, inclusive sobre os testes domiciliares, discriminação genética e privacidade sobre os resultados dos testes.

Resumo dos conceitos

- As restrições no estudo genético dos traços humanos incluem incapacidade para conduzir cruzamentos controlados, longo tempo de geração, família com poucos membros e dificuldade de separar as influências genéticas das ambientais. O heredograma é usado para estudar a herança dos traços nos seres humanos
- Os traços autossômicos recessivos aparecem com a mesma frequência em ambos os sexos e parecem pular gerações. Quando os pais são heterozigotos para um traço autossômico recessivo particular, aproximadamente 25% de seus descendentes terão o traço. É mais provável que os traços recessivos surjam em famílias com consanguinidade (acasalamento de parentes próximos)
- Os traços autossômicos dominantes aparecem com a mesma frequência em ambos os sexos e não pulam gerações. Quando um dos pais é afetado e heterozigoto para um traço autossômico dominante, aproximadamente metade de seus descendentes terá o traço. As pessoas não afetadas não transmitem um traço autossômico dominante para seus descendentes
- Os traços recessivos ligados ao X aparecem com maior frequência nos homens que nas mulheres. Quando uma mulher é portadora heterozigota de um traço recessivo ligado ao X e um homem não é afetado, aproximadamente metade de seus filhos terá o traço e metade de suas filhas será portadora não afetada. Os traços ligados ao X não são transmitidos do pai para o filho
- Os traços dominantes ligados ao X aparecem em homens e mulheres, embora sejam mais frequentes nas mulheres. Eles não pulam gerações. Os homens afetados transmitem um traço dominante ligado ao X para todas as suas filhas e nenhum dos seus filhos. As mulheres heterozigotas transmitem o traço para metade de seus filhos e metade de suas filhas
- Os traços ligados ao Y surgem apenas nos homens e são transmitidos do pai para todos os seus filhos
- Uma concordância maior de um traço nos gêmeos monozigóticos que nos dizigóticos indica influência genética no traço; concordância menor que 100% nos gêmeos monozigóticos indica influência ambiental no traço
- As semelhanças entre as crianças adotadas e seus pais biológicos indicam a importância dos fatores genéticos na expressão de um traço; as semelhanças entre as crianças adotadas e seus pais adotivos (sem relação genética) indicam a influência dos fatores ambientais
- O aconselhamento genético fornece informações e apoio a pessoas preocupadas sobre a hereditariedade nas suas famílias
- O exame genético inclui diagnóstico pré-natal, triagem para alelos responsáveis por doenças nos recém-nascidos, detecção de pessoas heterozigotas para alelos recessivos e exame pré-sintomático para a existência de um alelo responsável por doença em pessoas de risco
- Interpretar os testes genéticos pode ser complicado por haver múltiplas mutações, penetrância incompleta e influência dos fatores ambientais
- A disponibilidade de testes genéticos de coleta domiciliares tem levantado questões sobre a adequação das informações fornecidas, a ausência de aconselhamento genético, exatidão, privacidade e usos práticos de alguns testes

- O exame genético tem levantado questões sobre a discriminação genética e a privacidade dos resultados dos testes. O Genetic Information Nondiscrimination Act proíbe o uso das informações genéticas para decidir sobre seguro saúde e emprego.

Termos importantes

Aconselhamento genético
Amniocentese
Amostragem das vilosidades coriônicas
Cariótipo
Concordância
Consanguinidade
Diagnóstico genético pré-implante (PGD)
Diagnóstico genético pré-natal não invasivo
Exame genético pré-sintomático
Gêmeos dizigóticos
Gêmeos monozigóticos
Genetic Information Nondiscrimination Act (GINA)
Heredograma
Probando
Seleção de célula fetal
Teste de triagem no sangue materno
Teste genético de coleta domiciliar
Triagem de heterozigoto
Triagem do recém-nascido
Ultrassonografia

Respostas da Checagem dos conceitos

1. b.
2. Ele pode pular gerações quando uma nova mutação surge ou o traço tem penetrância reduzida.
3. Se recessivo ligado ao X, o traço não será transmitido do pai para o filho.
4. c.
5. Se o traço for ligado ao Y, um homem afetado transmitiria para todos os seus filhos, enquanto, se o traço for autossômico e limitado pelo sexo, os homens heterozigotos afetados o transmitiriam apenas para metade dos seus filhos, em média.
6. d.
7. c.
8. d.
9. O diagnóstico genético pré-implante determina a existência de genes causadores de doenças em um embrião no estágio inicial antes de ele ser implantado no útero e iniciar a gestação. O diagnóstico genético pré-natal determina a existência de genes ou cromossomos causadores de doenças em um feto em desenvolvimento.

Problemas desenvolvidos

Problema 1

Joana tem os “dedos das mãos curtos” (braquidactilia). Ela tem dois irmãos mais velhos que são gêmeos idênticos e ambos têm dedos das mãos curtos. Ela tem duas irmãs mais jovens com dedos das mãos normais. Sua mãe tem dedos das mãos normais e seu pai tem dedos das mãos curtos. Sua avó paterna tem dedos curtos, seu avô paterno, que é falecido, tinha dedos normais. Os avôs maternos têm dedos normais. Joana se casou com Tom que tem dedos normais, eles adotaram um filho chamado Bill que tem dedos normais. Os pais biológicos de Bill têm dedos normais. Após adotarem Bill, Joana e Tom tiveram duas crianças: uma filha mais velha com dedos curtos e um filho mais jovem com dedos normais.

a. Usando símbolos e marcações, desenhe o heredograma para ilustrar a herança dos dedos das mãos curtos na família de Joana.
b. Qual é a forma mais provável de herança para os dedos das mãos curtos nessa família?
c. Se Joana e Tom tiverem outro filho biológico, qual é a probabilidade (com base na sua resposta para *b*) de que essa

criança tenha dedos curtos?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

a. Um heredograma para representar a família, desenhe com os símbolos e marcações corretos.
b. A forma mais provável de herança para os dedos curtos.
c. A probabilidade de que a próxima criança de Joana e Tom tenha dedos das mãos curtos.

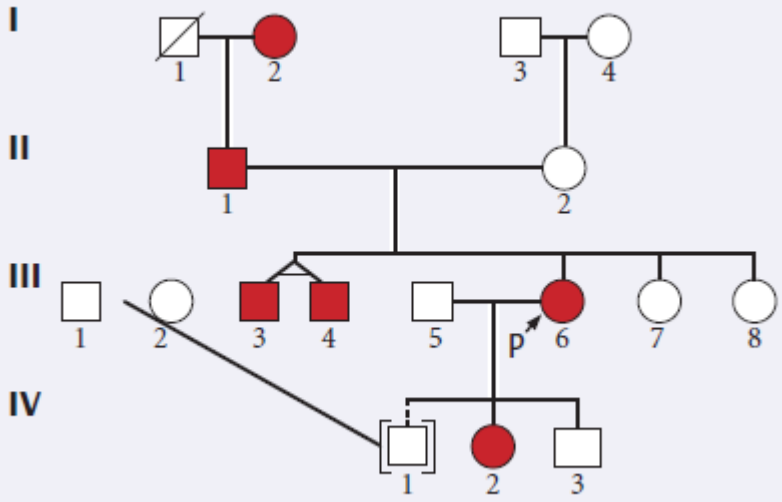
Quais informações são fornecidas para solucionar o problema?
Os fenótipos de Joana e Tom e dos membros de sua família.

Para a solução deste problema, revise:
Informações sobre heredogramas, na Seção 6.2.

Etapas para a solução

a. No heredograma da família, identifique as pessoas com o traço (dedos das mãos curtos) com círculos cheios (mulheres) e quadrados cheios (homens). Conecte os irmãos gêmeos idênticos de Joana à linha acima ao desenhar linhas diagonais com uma linha horizontal entre elas. Inclua a criança adotada por Joana e Tom entre parênteses, conecte-o aos seus pais biológicos ao desenhar uma linha diagonal e a seus pais adotivos com uma linha pontilhada.

Dica: Veja o Quadro 6.2 para uma revisão dos símbolos usados em um heredograma.



b. A forma mais provável de herança para os dedos das mãos curtos nessa família é a autossômica dominante. O traço aparece com a mesma frequência em homens e mulheres e não pula gerações. Quando um genitor tem o traço, ele surge em aproximadamente metade dos seus filhos e filhas, embora o número de crianças na família seja pequeno. Podemos eliminar a herança ligada ao Y porque o traço é encontrado nas mulheres assim como nos homens. Se os dedos curtos fossem um traço recessivo ligado ao X, seria esperado que as mulheres com o traço transmitissem para seus filhos, mas Joana (III-6), que tem dedos curtos, teve um filho com dedos normais. Para os traços dominantes ligados ao X, os homens afetados transmitiriam o traço para todas suas filhas; como o homem II-1 tem dedos curtos e teve duas filhas sem dedos curtos (III-7 e III-8), sabemos que o traço não pode ser dominante ligado ao X. É improvável que o traço seja autossômico recessivo, porque ele não pula gerações e aproximadamente metade dos filhos de pessoas afetadas tem o traço.

Dica: Veja o Quadro 6.3 para revisar as características das diferentes formas de herança.

c. Se ter os dedos curtos é autossômico dominante, Tom deve ser homozigoto (*bb*) pois ele tem dedos normais. Joana tem de ser heterozigota (*Bb*) pois ela e Tom tiveram filhos com dedos curtos e dedos normais. Em um cruzamento entre

um heterozigoto e homozigoto, espera-se que metade dos filhos seja heterozigota e a outra metade seja homozigota (*Bb* × *bb* → 1/2 *Bb*, 1/2 *bb*); então a probabilidade de que a próxima criança de Joana e Tom tenha dedos curtos é 50%.

Problema 2

Os valores de concordância para vários traços são medidos nos gêmeos homozigóticos e gêmeos dizigóticos, e os resultados são apresentados no quadro a seguir. Para cada traço, indique se as taxas de concordância sugerem influências genéticas, influências ambientais ou ambas. Explique seu raciocínio.

	Concordância (%)	
Característica	**Monozigótico**	**Dizigótico**
a. Tipo sanguíneo AB0	100	65
b. Diabetes melito	85	36
c. Ingestão de café	80	80
d. Tabagismo	75	42
e. Esquizofrenia	53	16

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Para cada traço, se ele é influenciado por fatores genéticos, fatores ambientais ou ambos.

Quais informações são fornecidas para solucionar o problema?
Valores de concordância dos gêmeos monozigóticos e dizigóticos para cada traço.

Para a solução deste problema, revise:
Concordância em gêmeos, na Seção 6.3.

Etapas para a solução

a. A concordância para o tipo sanguíneo AB0 nos gêmeos monozigóticos é 100%. Essa alta concordância nos gêmeos monozigóticos não indica por si só uma base genética para o traço. Como a concordância é muito menor nos gêmeos dizigóticos, podemos assegurar que os genes influenciam a determinação das diferenças nos tipos sanguíneos AB0.

> Lembrete: A influência dos fatores genéticos no traço é indicada pela maior concordância nos gêmeos monozigóticos do que nos gêmeos dizigóticos.

b. A concordância para diabetes melito é muito maior nos gêmeos monozigóticos do que nos gêmeos dizigóticos, portanto, podemos concluir que os fatores genéticos têm algum papel na suscetibilidade a diabetes. O fato de que os gêmeos monozigóticos mostram concordância menor que 100% sugere que os fatores ambientais também têm um papel.

> Lembrete: Concordância menor que 100% para gêmeos monozigóticos indica que os fatores ambientais influenciam o traço.

c. Os gêmeos monozigóticos e dizigóticos exibem a mesma concordância para beber café, então podemos concluir que existe pouca influência genética nessa característica. O fato de que os gêmeos monozigóticos mostram concordância menor que 100% sugere que os fatores ambientais influenciam o traço.

d. A concordância para tabagismo é menor nos gêmeos dizigóticos que nos gêmeos monozigóticos, então

parece que os fatores genéticos influenciam a tendência. O fato de que os gêmeos monozigóticos mostram concordância menor que 100% sugere que os fatores ambientais também influenciam o traço.

e. Os gêmeos monozigóticos exibem concordância muito maior para a esquizofrenia que os gêmeos dizigóticos, então podemos concluir que os fatores genéticos influenciam o transtorno psiquiátrico. Como a concordância dos gêmeos monozigóticos é muito menor que 100%, também podemos concluir que fatores ambientais também influenciam o traço.

Questões de compreensão

Seção 6.1

1. Quais são os três fatores que complicam o estudo da herança das características humanas?

Seção 6.2

2. Quem é o probando em um heredograma? O probando é sempre encontrado na última geração do heredograma? Por que sim ou não?
3. Para cada uma das seguintes formas de herança, descreva as características que serão exibidas em um heredograma no qual existe o traço: herança autossômica recessiva, autossômica dominante, recessiva ligada ao X, dominante ligada ao X, ligada ao Y.
4. Qual é a diferença entre o heredograma de um traço autossômico recessivo para o heredograma de um traço recessivo ligado ao X?
5. Além do fato de que um traço ligado ao Y surge apenas nos homens, qual é a diferença entre o heredograma de um traço ligado ao Y para o heredograma de um traço autossômico dominante?

Seção 6.3

6. Quais são os dois tipos de gêmeos que existem e como eles surgem?
7. Explique como uma comparação de concordância nos gêmeos monozigóticos e dizigóticos pode ser usada para determinar quanto da expressão de um traço é influenciada pelos genes ou pelos fatores ambientais.
8. Como os estudos sobre adoção são usados para separar os efeitos dos genes e do ambiente no estudo das características humanas?

Seção 6.4

9. O que é aconselhamento genético?
*10. Indique pelo menos quatro diferentes motivos para procurar o aconselhamento genético.
11. Defina em poucas palavras triagem de recém-nascidos, triagem de heterozigoto, exame pré-sintomático e diagnóstico pré-natal.
*12. Compare as vantagens e desvantagens da amniocentese *versus* amostragem das vilosidades coriônicas para o diagnóstico pré-natal.
13. O que é diagnóstico genético pré-implante?
14. Qual é a diferença entre a triagem de heterozigoto e o exame genético pré-sintomático?
*15. O que são os testes genéticos de coleta domiciliares? Quais são as questões envolvidas com esse tipo de teste?
*16. Quais são as atividades que o Genetic Information Nondiscrimination Act proíbe?
17. Como o exame genético poderia levar à discriminação genética?

Questões e problemas aplicados

Seção 6.1

*18. Se os seres humanos têm características que os tornam inadequados para a análise genética, como longo tempo de geração, família pequena e acasalamentos não controlados, por que os geneticistas persistem no estudo? Indique vários motivos por que os seres humanos são o foco de tantos estudos genéticos.

Seção 6.2

*19. Joe é daltônico. Seus pais têm visão colorida normal, mas seu avô materno é daltônico. Os outros avós de Joe têm visão colorida normal. Ele tem três irmãs, Patty, Betsy e Lora, todas com visão normal. Sua irmã mais velha, Patty, é casada

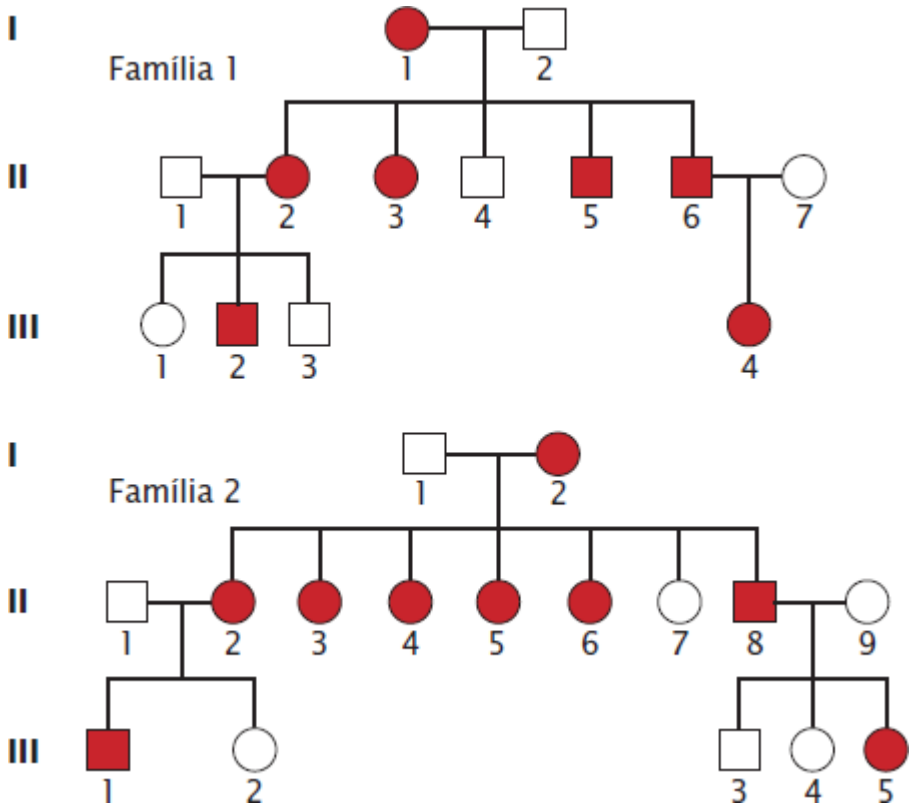
com um homem com visão normal, eles têm duas crianças: um menino de 9 anos daltônico e uma menina de 4 anos com visão normal.

a. Usando os símbolos e rótulos, desenhe um heredograma da família de Joe.
b. Qual é a forma mais provável de herança para o daltonismo na família de Joe?
c. Se ele se casar com uma mulher com história familiar de daltonismo, qual é a probabilidade de que sua primeira criança seja um menino daltônico?
d. Se ele se casar com uma mulher portadora do alelo para daltonismo, qual é a probabilidade de que sua primeira criança seja um menino daltônico?
e. Se Patty e seu marido tiverem outra criança, qual é a probabilidade de essa criança ser um menino daltônico?

20. Considere o heredograma apresentado na Figura 6.3.
    a. Se a pessoa IV-7 se casar com uma pessoa que não seja afetada pela síndrome de Waardenburg, qual é a probabilidade de que sua primeira criança tenha essa síndrome? Explique seu raciocínio.
    b. Se as pessoas IV-4 e IV-5 se casarem e tiverem uma criança, qual é a probabilidade de que essa criança tenha a síndrome de Waardenburg? Explique seu raciocínio.

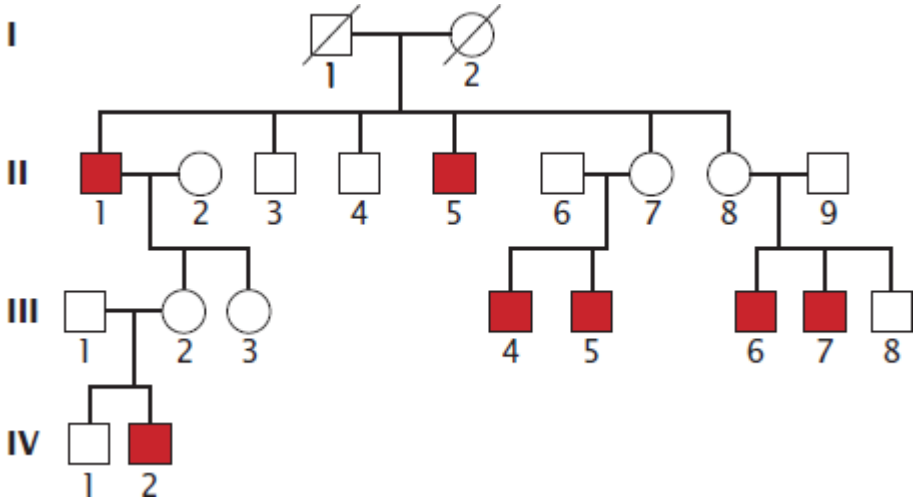
ANÁLISE DE DADOS

21. Muitos estudos sugerem uma forte predisposição genética a enxaquecas, mas a forma de herança não está clara. L. Russo *et al*. examinaram enxaquecas em várias famílias, duas das quais são apresentadas a seguir (L. Russo *et al*. 2005. *American Journal of Human Genetics* 76:327-333). Qual é a forma mais provável de herança para a enxaqueca nessas famílias? Explique seu raciocínio.



ANÁLISE DE DADOS

*22. A doença de Dent é um distúrbio raro dos rins, no qual a reabsorção do soluto filtrado está prejudicada e ocorre progressiva insuficiência renal. R. R. Hoopes *et al*. estudaram as mutações associadas à doença de Dent na família a seguir (R. R. Hoopes *et al*. 2005. *American Journal of Human Genetics* 76:260-267).

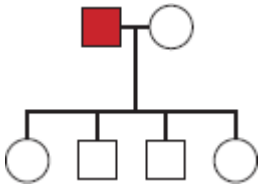


a. Com base nesse heredograma, qual é o modo mais provável de herança para a doença? Explique seu raciocínio.

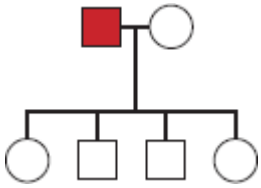
b. A partir da sua resposta para *a*, indique os genótipos mais prováveis para todas as pessoas no heredograma.

23. Um homem com um traço genético específico raro se casa com uma mulher não afetada e eles têm quatro crianças. Os heredogramas dessa família são apresentados nas partes *a* a *e*, mas a existência ou ausência do traço na criança não está indicada. Para cada tipo de herança, indique quantas crianças de cada sexo devem expressar o traço ao preencher os círculos e quadrados adequados. Considere que o traço é raro e totalmente penetrante.

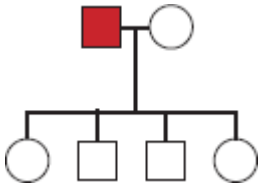
a. Traço autossômico recessivo.



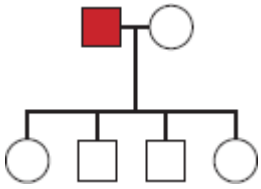
b. Traço autossômico dominante.



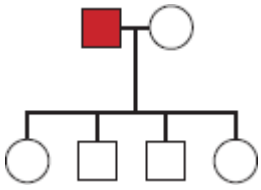
c. Traço recessivo ligado ao X.



d. Traço dominante ligado ao X.



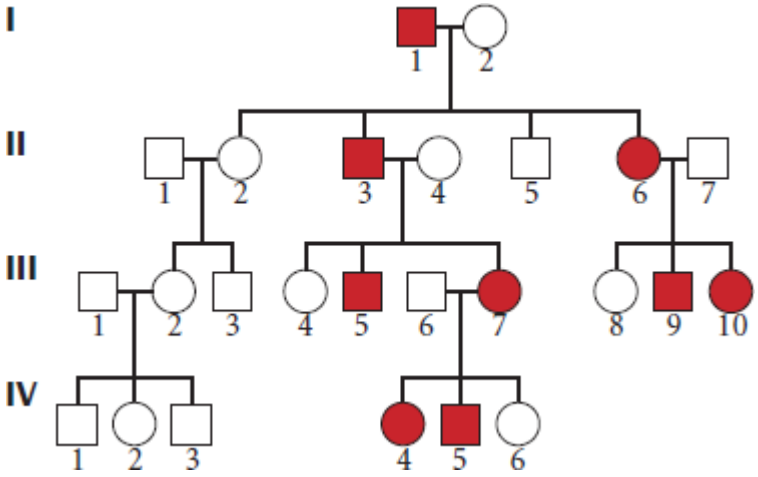
e. Traço ligado ao Y.



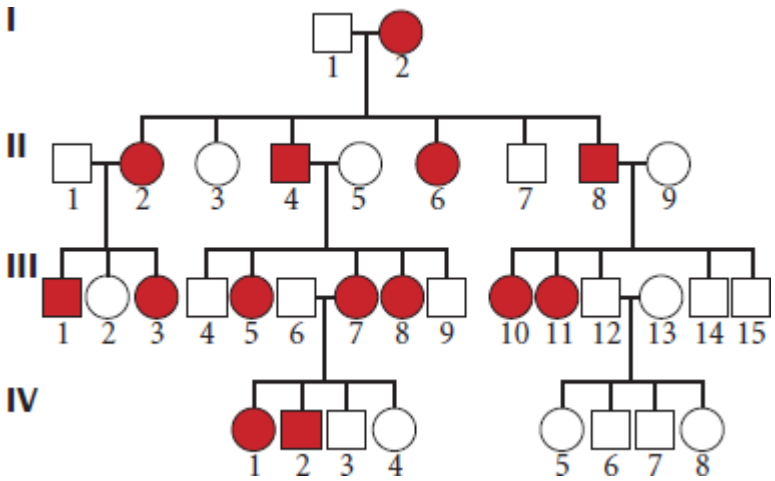
*24. Para cada um dos heredogramas a seguir, indique o modo mais provável de herança, considerando que o traço é raro.

Explique seu raciocínio com detalhes.

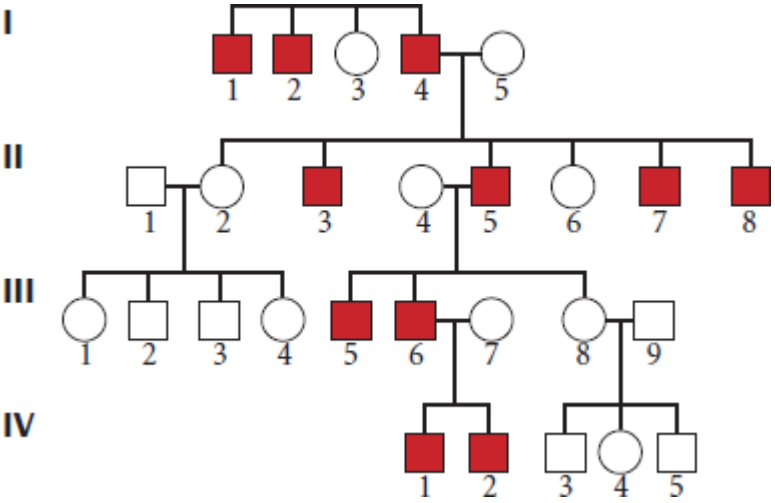
a.



b.

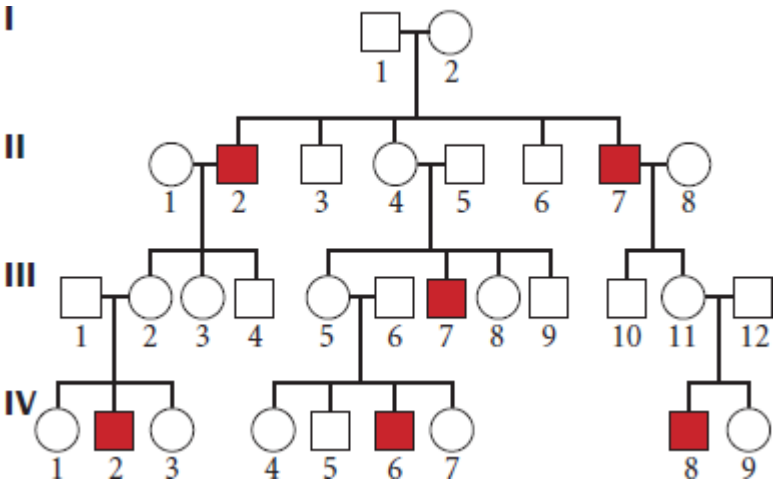


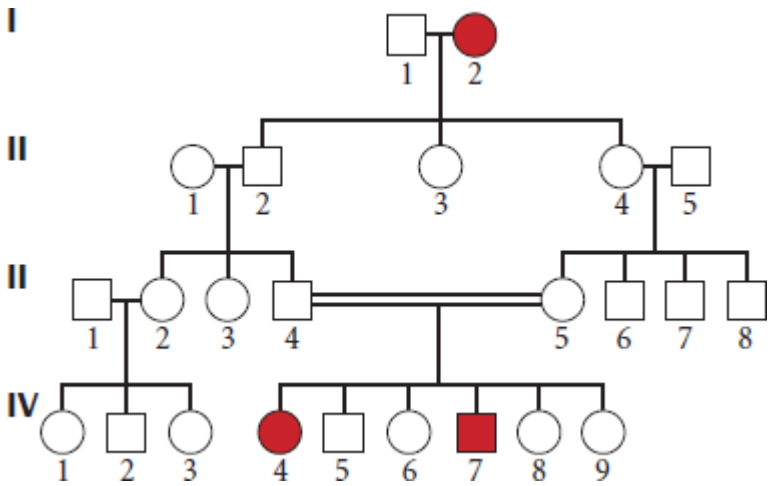
c.



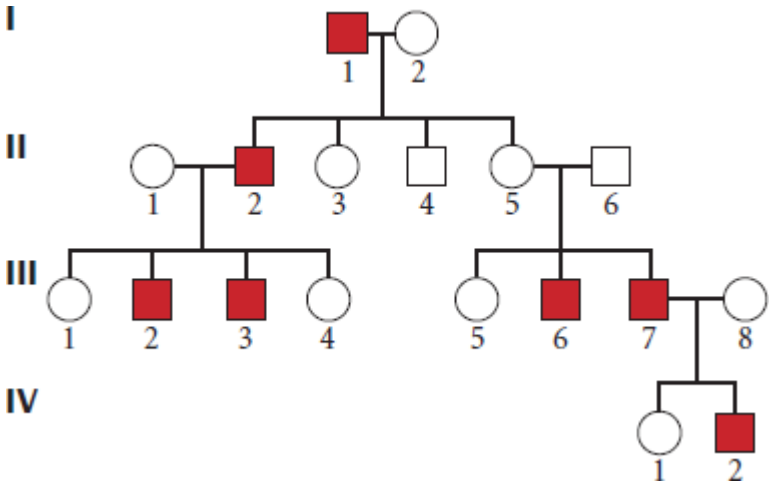
d.

e.

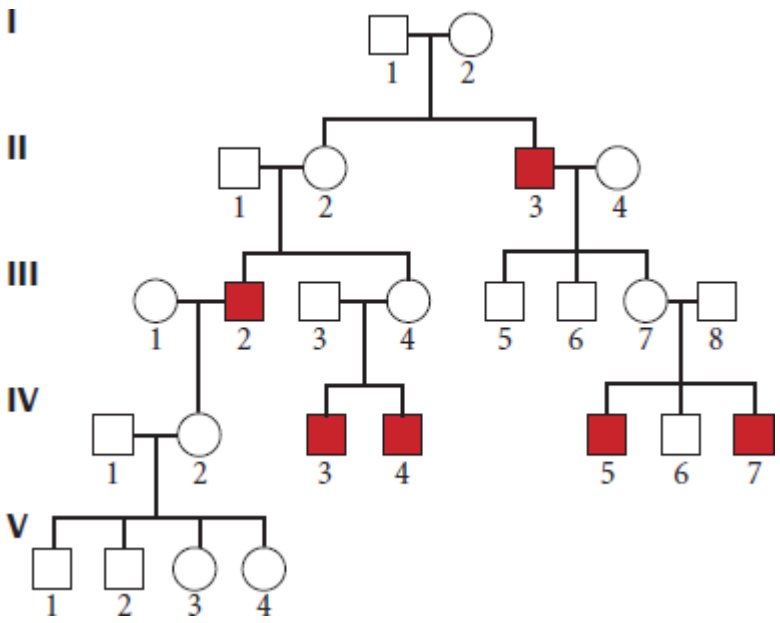




25. O traço representado no heredograma a seguir é expresso apenas nos homens da família. O traço está ligado ao Y? Por que sim ou não? Se você acredita que o traço não está ligado ao Y, proponha outra explicação para sua herança.



26. O heredograma a seguir ilustra a herança da síndrome de Nance-Horan, uma rara condição genética na qual as pessoas afetadas têm catarata e dentes com formato anormal.

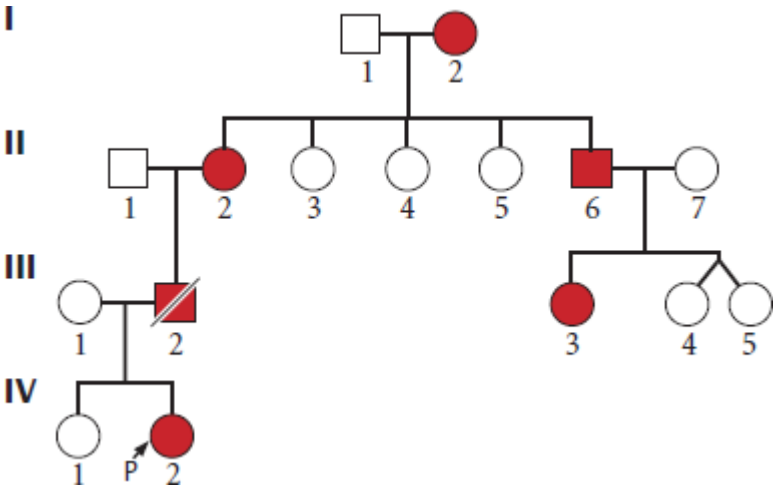


(Heredograma adaptado de D. Stambolian, R. A. Lewis, K. Buetow, A. Bond, and R. Nussbaum. 1990. American Journal of Human Genetics 47:15.)

a. Com base neste heredograma, qual é o modo mais provável de herança para essa síndrome?
b. Se o casal III-7 e III-8 tiver outra criança, qual é a probabilidade de que ela tenha a síndrome de Nance-Horan?
c. Se III-2 e III-7 acasalarem, qual é a probabilidade de uma de suas crianças ter essa síndrome?

ANÁLISE DE DADOS

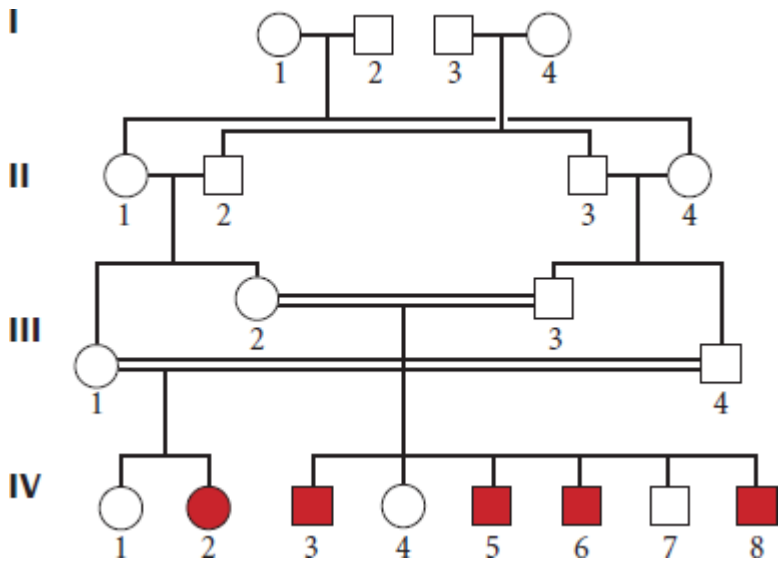
27. O heredograma a seguir ilustra a herança do cabelo anelar, uma condição na qual cada pelo é diferenciado em zonas claras e escuras. Qual modo ou quais modos de herança são possíveis para o traço de cabelo anelar nessa família?



(Heredograma adaptado de L. M. Ashley and R. S. Jacques. 1950. Journal of Heredity 41:83.)

ANÁLISE DE DADOS

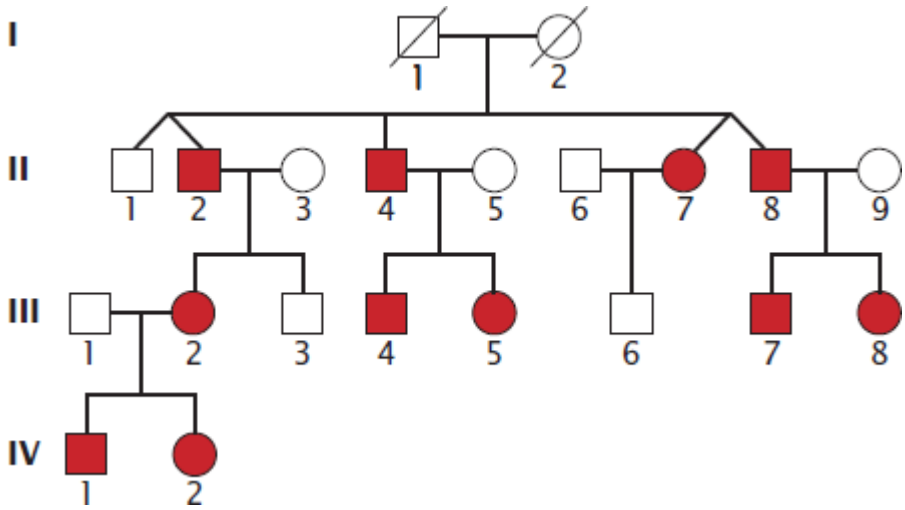
28. A ectrodactilia é uma rara condição na qual os dedos estão ausentes e a mão é dividida. Em geral essa condição é herdada como um traço autossômico dominante. Ademar Freire-Maia relatou o surgimento de ectrodactilia em uma família em São Paulo, Brasil, cujo heredograma está apresentado aqui. Esse heredograma está de acordo com a herança autossômica dominante? Se não, qual é o modo de herança mais provável? Explique seu raciocínio.



(Heredograma adaptado de A. Freire-Maia. 1971. Journal of Heredity 62:53.)

ANÁLISE DE DADOS

29. A ausência completa de um ou mais dentes (agenesia dentária, AD) é um traço comum nos seres humanos; de fato, a análise mostra que mais de 20% dos seres humanos não têm um ou mais de seus terceiros molares. Entretanto, a ausência mais grave de dentes, definida como a ausência de seis ou mais dentes, é menos comum e frequentemente é uma condição hereditária. L. Lammi *et al*. examinaram a agenesia dentária na família finlandesa apresentada no heredograma abaixo (L. Lammi 2004. *American Journal of Human Genetics* 74:1043-1050).



(Heredograma adaptado de L. Lammi 2004. American Journal of Human Genetics 74:1043-1050.)

a. Qual é a forma mais provável de herança para a agenesia dentária nessa família? Explique seu raciocínio.
b. Os dois conjuntos de gêmeos nessa família são monozigóticos ou dizigóticos? Qual é o fundamento da sua resposta?
c. Se IV-2 se casar com um homem com todos os dentes, qual é a probabilidade de que seu filho tenha agenesia dentária?
d. Se as pessoas III-2 e III-7 se casarem e tiverem um filho, qual é a probabilidade de que essa criança tenha a condição?

Seção 6.3

*30. Um geneticista estuda várias características em gêmeos monozigóticos e dizigóticos, obtendo as concordâncias listadas a seguir. Para cada característica, indique se as taxas de concordância sugerem influências genéticas, influências ambientais ou ambas. Explique seu raciocínio.

	Concordância (%)	
Característica	**Monozigótico**	**Dizigótico**
Enxaqueca	60	30

Cor dos olhos	100	40
Sarampo	90	90
Pé torto	30	10
Hipertensão arterial	70	40
Destreza	70	70
Tuberculose	5	5

31. Com base nas taxas de concordância apresentadas no Quadro 6.2, a variação na artrite reumatoide é influenciada por fatores genéticos, ambientais ou ambas? Explique seu raciocínio.

32. M. T. Tsuang *et al*. estudaram a dependência química em pares de gêmeos do sexo masculino (M. T. Tsuang *et al*. 1996. *American Journal of Medical Genetics* 67:473-477). Eles descobriram que 82 de 313 gêmeos monozigóticos eram concordantes para o abuso de uma ou mais drogas ilícitas, enquanto 40 de 243 gêmeos dizigóticos eram concordantes para o mesmo traço. Calcule as taxas de concordância para abuso de drogas nesses gêmeos. Com base nos dados encontrados, que conclusão você pode tirar sobre o papel dos fatores genéticos e ambientais no abuso de drogas?

*33. No estudo da esquizofrenia (um transtorno mental que inclui desorganização do pensamento e fuga da realidade), os pesquisadores estudaram a sua prevalência nos pais biológicos e adotivos de pessoas que foram adotadas quando crianças, encontrando os seguintes resultados:

	Prevalência da esquizofrenia	
Pessoas adotadas	**Pais biológicos**	**Pais adotivos**
Com esquizofrenia	12	2
Sem esquizofrenia	6	4

(Fonte: S. S. Kety et al. 1978. in The Nature of Schizophrenia: New Approaches to Research and Treatment, L. C. Wynne, R. L. Cromwell, and S. Matthysse, Eds. New York: Wiley, 1978, pp. 25-37.)

Que conclusão sobre o papel da genética na esquizofrenia você pode tirar a partir desses resultados? Explique seu raciocínio.

34. Que conclusões são apoiadas pela Figura 6.11?
    a. Os pais adotivos de crianças obesas têm um IMC maior que os pais adotivos das crianças magras.
    b. As mães adotivas de crianças magras têm um IMC menor que as mães adotivas das crianças obesas.
    c. Os pais biológicos de crianças obesas têm um IMC maior que os pais adotivos das crianças magras.
    d. a e b.
    e. a e c.

Seção 6.4

35. Quais (se alguma) questões éticas podem surgir com o uso disseminado do diagnóstico genético pré-natal não invasivo, que pode ser feito muito antes da amniocentese ou da amostragem das vilosidades coriônicas?

Questões desafiadoras

Seção 6.1

36. Muitos estudos genéticos, sobretudo os de traços recessivos, concentram-se em pequenas populações humanas isoladas, como as de ilhas. Sugira uma ou mais vantagens que as populações isoladas podem ter para o estudo dos traços recessivos.

Seção 6.2

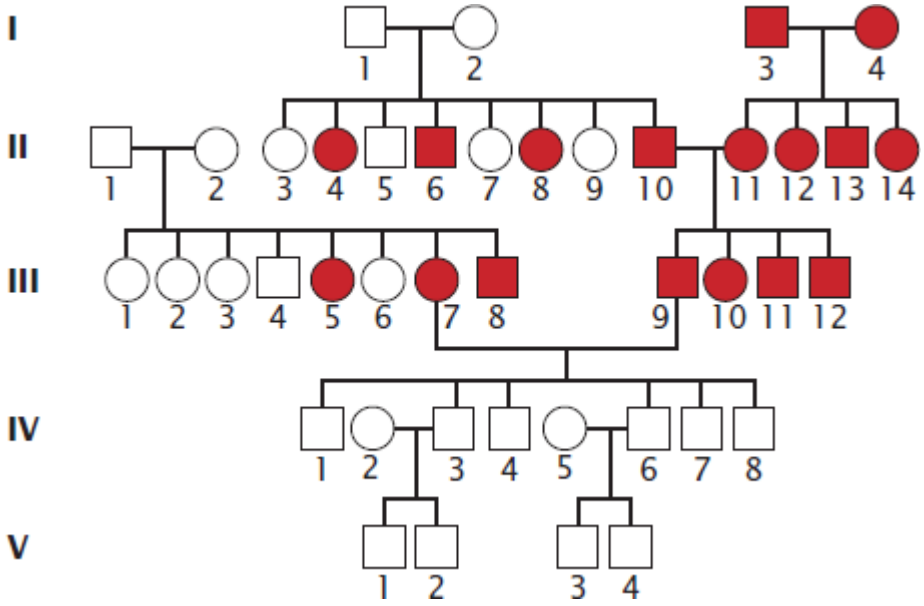
37. Desenhe um heredograma que represente um traço autossômico dominante, limitado pelo sexo nos homens e que exclua a possibilidade de que o traço esteja ligado ao Y.

ANÁLISE DE DADOS

38. A. C. Stevenson e E. A. Cheeseman estudaram a surdez em uma família na Irlanda do Norte e registraram o heredograma ao lado (A. C. Stevenson and E. A. Cheeseman. 1956. *Annals of Human Genetics* 20:177-231).
    a. Se você considerar apenas as gerações I a III, qual é o modo mais provável de herança para esse tipo de surdez?
    b. Forneça uma explicação possível para o cruzamento entre III-7 e III-9 e os resultados para as gerações IV e V.

Seção 6.3

39. A geração de gêmeos dizigóticos tem, com frequência, caráter familiar, e sua frequência varia entre grupos étnicos, enquanto a geração de gêmeos monozigóticos raramente é familiar e sua frequência é constante em grupos étnicos. Essas observações foram interpretadas como evidências de base genética sobre a variação na geração de gêmeos dizigóticos, mas pouca base genética para a variação nos gêmeos monozigóticos. Você pode sugerir um possível motivo para essas diferenças nas tendências genéticas na geração de gêmeos dizigóticos e monozigóticos?



(Heredograma adaptado de A. C. Stevenson and E. A. Cheeseman. 1956. Annals of Human Genetics 20:177-231.)

7

Ligação Gênica, Recombinação e Mapeamento de Genes Eucarióticos

Os genes ligados e a calvície

Para muitos, a calvície é a maldição da masculinidade. Vinte e cinco por cento dos homens começam a perder cabelo antes dos 30 anos, e quase metade apresenta algum grau de calvície até os 50 anos. Nos EUA, ela afeta mais de 40 milhões de homens, e centenas de milhares de dólares são gastos todos os anos no tratamento da perda de cabelo. Não é apenas uma questão de vaidade: os homens calvos apresentam maior propensão para doenças cardíacas, hipertensão arterial e câncer de próstata.

A calvície pode surgir por diferentes motivos, incluindo doença, lesão, medicamentos e hereditariedade. O tipo mais comum de calvície observada nos homens é a calvície de padrão masculino – tecnicamente conhecida como alopecia androgenética –, na qual o cabelo é perdido de forma prematura a partir da frente e do topo da cabeça. Mais de 95% da perda de cabelo nos homens é desse tipo. Embora também seja observada nas mulheres, em geral ela é expressa de maneira mais sutil, como leve afinamento do cabelo. O traço é estimulado pelos hormônios sexuais masculinos (andrógenos), como é confirmado pela observação de que os homens castrados jovens raramente ficam calvos (embora isso não seja recomendado como tratamento preventivo).

Foi identificada forte influência hereditária nesse padrão de alopecia, mas o modo exato de herança é controverso. Um estudo inicial sugeriu que o traço é autossômico dominante nos homens e recessivo nas mulheres, um exemplo de traço influenciado pelo sexo (ver *Capítulo 5*). Outras evidências e a cultura popular comum sugerem que um homem herda a calvície do lado materno da família, ao exibir uma herança ligada ao X.

Em 2005, o geneticista Axel Hillmer e seus colegas começaram a localizar o gene responsável pela alopecia androgenética. Eles suspeitaram que o gene pudesse estar localizado no cromossomo X, mas não conheciam sua exata localização. Para identificar a localização do gene, eles realizaram um estudo de análise de ligação, na qual eles procuraram por uma associação entre a herança da alopecia androgenética e a herança de variantes genéticas conhecidas por estarem localizadas no cromossomo X. As variantes genéticas usadas no estudo foram os polimorfismos de único nucleotídio (SNPs; do inglês, *single-nucleotide polymorphisms*), que são posições no genoma onde as pessoas apresentam variação em um único nucleotídio. Os geneticistas estudaram a herança da calvície-padrão e os SNPs em 95 famílias nas quais pelos menos dois irmãos desenvolveram esse tipo de calvície ainda jovens.

Os pesquisadores descobriram que a alopecia androgenética e os SNPs a partir do cromossomo X não eram herdados de forma independente, como previsto pelo princípio de Mendel de segregação independente. Em vez disso, eles eram herdados juntos, o que ocorre quando os genes estão fisicamente ligados no mesmo cromossomo e se separam juntos na meiose.

Como aprenderemos neste capítulo, a ligação entre os genes é rompida por um processo chamado de recombinação, ou *crossing over*, e a quantidade de recombinação entre os genes está diretamente relacionada com a distância entre eles. Em 1911, Thomas Hunt Morgan e seu aluno Alfred Sturtevant demonstraram nas moscas-da-fruta que os genes podem ser mapeados ao determinar as taxas de recombinação entre os genes. Ao usar o método nas famílias com a alopecia androgenética, Hillmer e seus colegas demonstraram que o gene para esse tipo de calvície está intimamente ligado a SNPs localizados na posição p12-22 no cromossomo X. Essa região inclui o gene do receptor de andrógenos, que codifica uma proteína que se liga aos hormônios sexuais masculinos. Considerando o claro envolvimento dos hormônios masculinos no desenvolvimento da calvície-padrão, é provável que o gene do receptor de andrógenos seja um candidato a agente causal desse traço. Outra análise revelou que certos alelos do gene do receptor de andrógenos estavam intimamente associados à herança da alopecia androgenética, e que é quase certo que o gene desse receptor era responsável por boa parte das diferenças no traço observadas nas famílias examinadas. Estudos adicionais realizados em 2008 descobriram que os genes nos cromossomos 3 e 20 também parecem contribuir para a expressão da alopecia androgenética. **Resolva o Problema 13**

Este capítulo explora a herança dos genes localizados no mesmo cromossomo. Esses genes ligados não obedecem estritamente ao princípio de Mendel da segregação independente; em vez disso, eles tendem a ser herdados juntos. Esta tendência requer uma nova abordagem para compreender sua herança e prever os tipos de descendentes produzidos. Uma informação crítica necessária para prever os resultados desses cruzamentos é o arranjo dos genes nos cromossomos; para isso será necessário pensar na relação entre genes e cromossomos. Para compreender a herança dos genes ligados é preciso fazer a conexão conceitual entre os genótipos

em um cruzamento e o comportamento dos cromossomos na meiose.

Vamos começar nossa exploração sobre ligação ao comparar primeiro a herança de dois genes ligados com a herança de dois genes que se separam de forma independente. Vamos, então, examinar como a recombinação separa os genes ligados. Esse conhecimento sobre ligação e recombinação será usado para prever os resultados dos cruzamentos genéticos nos quais os genes estão ligados e para mapear genes. Posteriormente, vamos nos concentrar nos métodos físicos para determinar as localizações cromossômicas dos genes. A seção final examina a variação nas taxas de recombinação.

7.1 Os genes ligados não se separam de forma independente

O Capítulo 3 introduziu os princípios da segregação e segregação independente de Mendel. Vamos agora revisar dois conceitos importantes: o princípio da segregação e o princípio da segregação independente. O primeiro afirma que cada organismo diploide tem dois alelos em um *locus* que se separam na meiose, cada alelo indo para um gameta. O segundo princípio fornece mais informações sobre o processo de segregação: ele afirma que, no processo de separação, os dois alelos em um *locus* atuam de maneira independente dos alelos em outros *loci*.

A segregação independente dos alelos leva à *recombinação*, a seleção dos alelos em novas combinações. Considere um cruzamento entre indivíduos homozigotos para dois diferentes pares de alelos: *AA BB* × *aa bb* O primeiro genitor, *AA BB*, produz gametas com os alelos *A B* e o segundo genitor, *aa bb*, produz gametas com alelos *a b*, gerando os descendentes F_1 com genótipo *Aa Bb* (**Figura 7.1**). A recombinação significa que, quando um dos descendentes F_1 se reproduz, a combinação dos alelos nos seus gametas pode ser diferente das combinações nos gametas de seus genitores. Em outras palavras, F_1 pode produzir gametas com novas combinações dos alelos *A b* ou *a B*, além dos gametas dos genitores com *A B* ou *a b*.

Mendel produziu seus princípios de segregação e segregação independente ao observar os descendentes dos cruzamentos genéticos, mas ele não tinha ideia sobre que processos biológicos produziam esses fenômenos. Em 1903, Walter Sutton propôs uma base biológica para os princípios de Mendel, chamada de teoria cromossômica da herança, que afirma que os genes são encontrados nos cromossomos (ver *Capítulo 3*). Vamos reformular os dois princípios de Mendel em relação à teoria cromossômica da herança. O princípio da segregação afirma que um organismo diploide tem dois alelos para uma característica, e cada um está localizado na mesma posição, ou *locus*, em cada um dos dois cromossomos homólogos. Esses cromossomos se separam na meiose, com cada gameta recebendo um homólogo. O princípio da segregação independente afirma que, na meiose, cada par de cromossomos homólogos se separa de forma independente dos outros pares de homólogos. Com essa nova perspectiva, é fácil ver que o número de cromossomos na maioria dos organismos é limitado e que com certeza existem mais genes do que cromossomos, então alguns genes devem estar presentes no mesmo cromossomo e não devem se separar dessa forma. Os genes localizados próximos no mesmo cromossomo são chamados de **genes ligados** e pertencem ao mesmo **grupo de ligação**. Os genes ligados se deslocam juntos na meiose, consequentemente chegando ao mesmo destino (mesmo gameta), e não é esperado que se separem de forma independente.

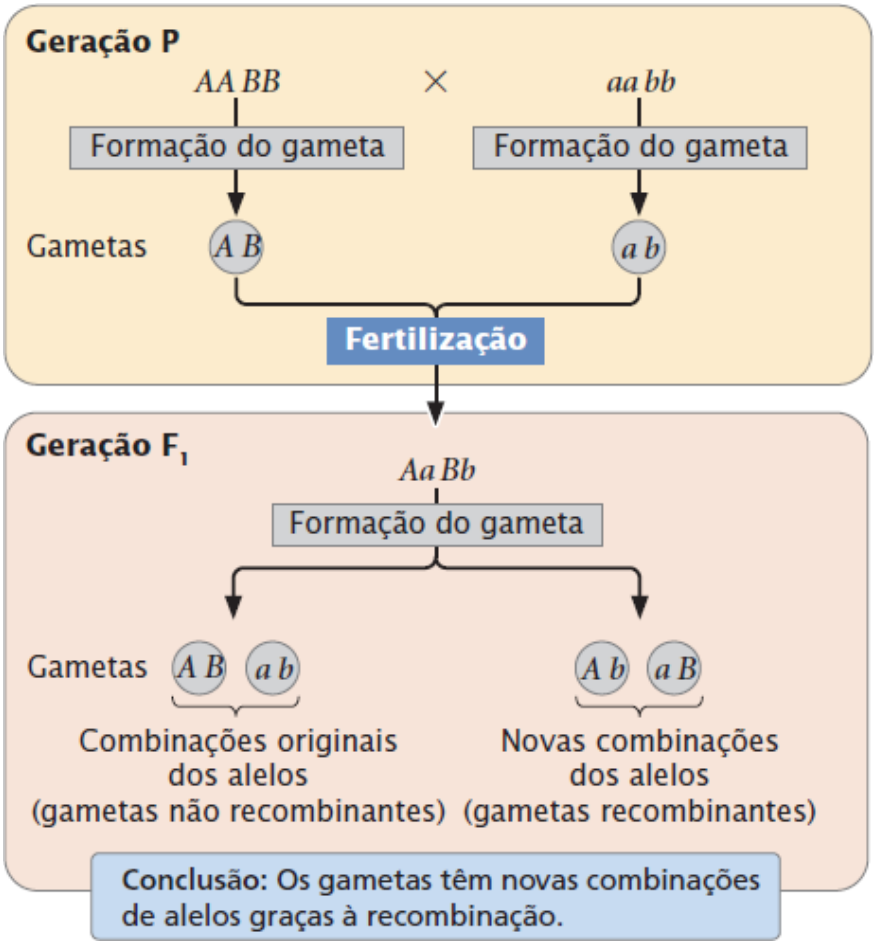


Figura 7.1 **A Recombinação é a seleção dos alelos em novas combinações.**

Todas as características examinadas por Mendel nas ervilhas apresentaram segregação independente e, após a redescoberta de seu trabalho, as primeiras características genéticas estudadas em outros organismos também se separavam da mesma forma. Como os genes poderiam ser carreados em um número limitado de cromossomos e mesmo assim se separarem de forma independente?

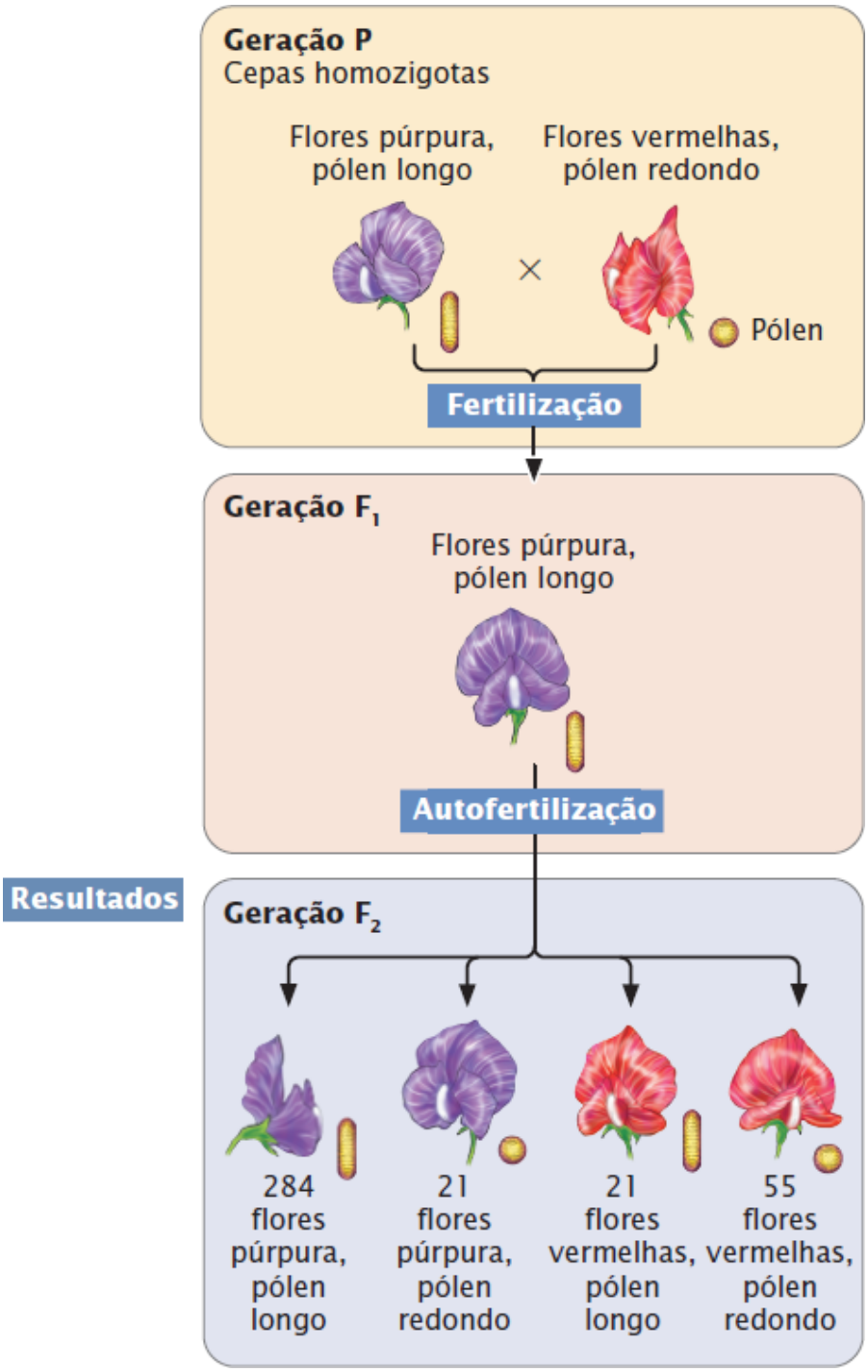
Essa aparente inconsistência entre o princípio da segregação independente e a teoria cromossômica da herança desapareceu assim que os biólogos começaram a encontrar características genéticas que não se separavam de forma independente. Um dos primeiros casos foi relatado nas ervilhas-de-cheiro por William Bateson, Edith Rebecca Saunders e Reginald C. Punnett, em 1905. Eles cruzaram uma cepa homozigota de ervilhas que tinha flores púrpura e grãos de pólen longos com uma cepa homozigota com flores vermelhas e grãos de pólen redondos. Todos os descendentes F_1 tinham flores púrpura e grãos de pólen longos, indicando que púrpura era dominante sobre vermelho e que longo era dominante sobre redondo. Quando eles intercruzaram F_1, os descendentes F_2 resultantes não surgiram na razão 9: 3: 3: 1 esperada com a segregação independente (**Figura 7.2**). Um excedente de plantas F_2 tinha flores púrpura e pólen longo ou flores vermelhas e pólen redondo (os fenótipos dos genitores). Embora Bateson, Saunder e Punnett não fossem capazes de explicar os resultados, agora sabemos que os dois *loci* que eles examinaram estavam próximos no mesmo cromossomo e, portanto, não se separam de forma independente.

Experimento

Pergunta: Os genes para a cor da flor e o formato do pólen nas ervilhas-de-cheiro se separam de forma independente?

Métodos

Cruze duas cepas homozigotas para dois traços diferentes.



Conclusão: Os descendentes F2 não surgem na razão

Figura 7.2 A separação não independente da cor da flor e formato do pólen na ervilha-de-cheiro.

7.2 Os genes ligados se separam juntos, enquanto o *crossing over* produz recombinação entre eles

Os genes que estão próximos no mesmo cromossomo geralmente se separam como uma unidade e são, portanto, herdados juntos. Às vezes, entretanto, os genes trocam de um cromossomo homólogo para outro por meio do *crossing over* (ver *Capítulo 2*), como ilustrado na **Figura 7.3**. O resultado do *crossing over* é a recombinação: ele rompe as associações dos genes que estão próximos no mesmo cromossomo. A ligação e o *crossing over* podem ser observados como processos que têm efeitos opostos; a ligação mantém determinados genes juntos, e o *crossing over* mistura esses genes, produzindo novas combinações. No Capítulo 5 estudamos várias exceções e extensões para os princípios de hereditariedade de Mendel. O conceito de genes ligados complica ainda mais as interpretações de resultados de cruzamentos genéticos. Entretanto, com a compreensão de como a ligação afeta a hereditariedade, podemos analisar os cruzamentos para os genes ligados e prever com sucesso os tipos de descendentes que serão produzidos.

Notação dos cruzamentos com ligação

Ao analisar os cruzamentos com genes ligados, não devemos saber apenas os genótipos dos indivíduos cruzados, mas também a disposição dos genes nos cromossomos. Para rastrear essa disposição, introduzimos um novo sistema de notação para os cruzamentos apresentados com os genes ligados. Considere um cruzamento entre um homozigoto para os alelos dominantes dos dois *loci* ligados e outro homozigoto para alelos recessivos nesses *loci* (*AA BB* × *aa bb*). É necessário escrever os alelos específicos como eles estão dispostos em cada um dos cromossomos homólogos para os genes ligados:

$$\frac{A \quad B}{A \quad B} \times \frac{a \quad b}{a \quad b}$$

Nesta notação, cada linha representa um dos dois cromossomos homólogos. Ao herdar um cromossomo de cada genitor, os descendentes F_1 terão o seguinte genótipo:

$$\frac{A \quad B}{a \quad b}$$

Está clara a importância de indicar os alelos em cada cromossomo neste caso. Um cromossomo tem dois alelos dominantes *A* e *B*, enquanto o cromossomo homólogo tem dois alelos recessivos *a* e *b*. A numeração pode ser simplificada pelo traçado de uma única linha, com a compreensão de que os genes localizados no mesmo lado da linha estão no mesmo cromossomo:

$$\frac{A \quad B}{a \quad b}$$

Esta notação pode ser ainda mais simplificada pela separação dos alelos em cada cromossomo por uma barra: *AB*/*ab*.

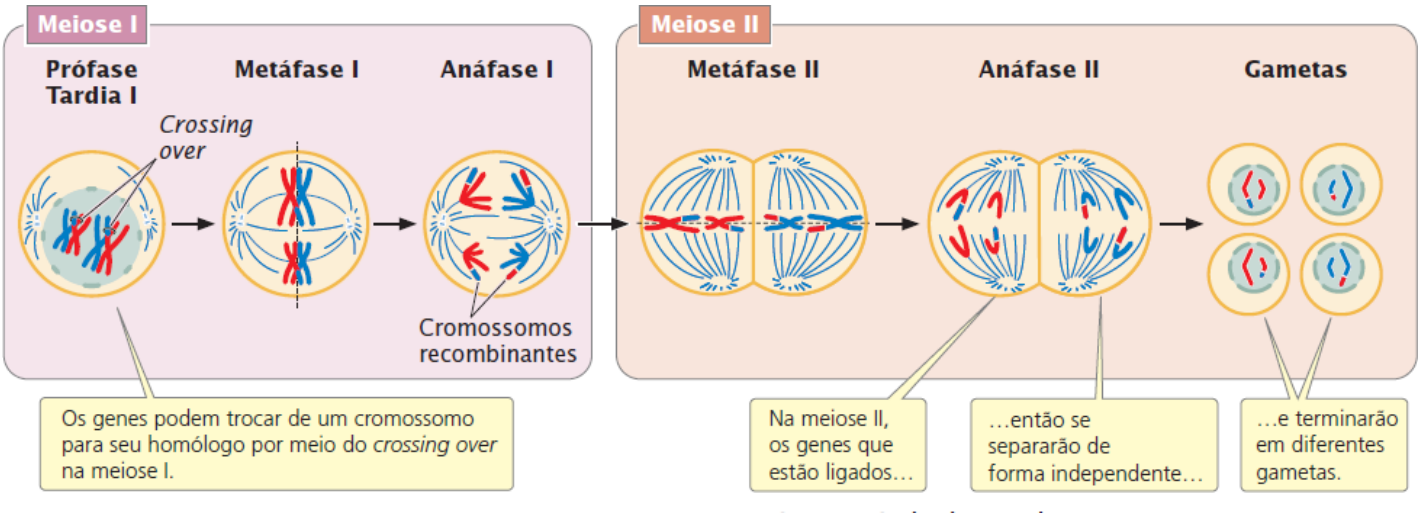


Figura 7.3 O *crossing over* ocorre na meiose e é responsável pela recombinação.

Lembre-se de que os dois alelos em um *locus* estão sempre localizados em diferentes cromossomos homólogos e, portanto, estão em lados opostos da linha. Consequentemente, *nunca* escreveríamos os genótipos como

$$\frac{A \quad a}{B \quad b}$$

porque os alelos *A* e *a* nunca podem estar no mesmo cromossomo. Também é importante manter a mesma ordem dos genes em ambos os lados da linha; assim *nunca* devemos escrever

$$\frac{A \quad B}{b \quad a}$$

porque isso indicaria que os alelos *A* e *b* são alélicos (estão no mesmo *locus*).

Ligação completa comparada com segregação independente

Vamos, a princípio, considerar o que acontece aos genes que exibem ligação completa: eles estão localizados muito próximos no mesmo cromossomo e não exibem *crossing over*. Em raros casos os genes estão completamente ligados, mas ao assumir que não

ocorre *crossing over*, podemos ver o efeito da ligação de forma mais clara. Então vamos considerar o que acontece quando os genes se separam de forma independente. Finalmente, vamos considerar os resultados obtidos se os genes estiverem ligados, mas exibirem um grau de *crossing over*.

Um cruzamento-teste revela os efeitos da ligação. Por exemplo, se for feito um cruzamento-teste de um indivíduo heterozigoto com um homozigoto recessivo (*Aa Bb* × *aa bb*), os alelos existentes nos gametas contribuídos pelo genitor heterozigoto serão expressos no fenótipo dos descendentes porque o genitor homozigoto não poderia contribuir com alelos dominantes, que poderiam mascará-los. Consequentemente, os traços que surgem nos descendentes revelam quais alelos foram transmitidos pelo genitor heterozigoto.

Considere um par de genes ligados nos tomates. Um dos genes afeta o tipo da folha: um alelo para folhas mosqueadas (*m*) é recessivo em relação ao alelo que produz folhas normais (*M*). Próximo, no mesmo cromossomo, o outro gene determina a altura da planta: um alelo para planta anã (*d*) é recessivo em relação ao alelo para planta alta (*D*).

O exame para a ligação pode ser feito com um cruzamento-teste, que requer uma planta heterozigota para ambas as características. Um geneticista pode produzir esta planta ao cruzar uma variedade de tomate homozigota para folhas normais e planta alta com uma variedade homozigota pra folhas mosqueadas e planta anã:

$$\text{P} \qquad \frac{M \quad D}{M \quad D} \times \frac{m \quad d}{m \quad d}$$

$$\downarrow$$

$$\text{F}_1 \qquad \frac{M \quad D}{m \quad d}$$

O geneticista, então, usaria os heterozigotos F_1 em um cruzamento-teste com plantas homozigotas para folhas mosqueadas e plantas anãs:

$$\frac{M \quad D}{m \quad d} \times \frac{m \quad d}{m \quad d}$$

Os resultados deste cruzamento estão representados na **Figura 7.4 A.** O heterozigoto produz dois tipos de gametas, alguns com os cromossomos *M D* e outros com o cromossomo *m d*. Como não ocorre *crossing over*, esses gametas são os únicos tipos produzidos pelo heterozigoto. Observe que eles têm apenas combinações de alelos existentes nos genitores originais: o alelo para folhas normais junto com o alelo para planta alta (*M* e *D*) ou o alelo para folhas mosqueadas junto com o alelo para planta anã (*m* e *d*). Os gametas que têm apenas as combinações originais dos alelos presentes nos genitores são **gametas não recombinantes**, ou gametas *parentais*.

O genitor homozigoto no cruzamento-teste produz apenas um tipo de gameta; ele tem o cromossomo *m d* e pares com um dos dois gametas gerados pelo genitor heterozigoto (ver Figura 7.4 A). Surgem dois tipos de descendentes: metade tem folhas normais e é alta:

$$\frac{M \quad D}{m \quad d}$$

e metade tem folhas mosqueadas e é anã:

$$\frac{m \quad d}{m \quad d}$$

Esses descendentes apresentam as combinações originais dos traços existentes na geração P e são **descendentes não recombinantes** ou descendentes *parentais*. Nenhuma das novas combinações dos dois traços, com folhas normais e planta anã ou folhas mosqueadas com planta alta aparece nos descendentes, porque os genes que afetam os dois traços estão completamente ligados e são herdados juntos. Novas combinações de traços podem surgir apenas se a conexão física entre *M* e *D* ou entre *m* e *d* for rompida.

Esses resultados são bem diferentes dos que são esperados quando os genes se separam de forma independente (**Figura 7.4 B**).

Se os *loci M* e *D* se separassem de forma independente, a planta heterozigota (*Mm Dd*) produziria quatro tipos de gametas: dois gametas não recombinantes com as combinações originais dos alelos (*M D* e *m d*) e dois com novas combinações de alelos (*M d* e *m D*). Os gametas com novas combinações de alelos são chamados de **gametas recombinantes**. Na segregação independente, os gametas não recombinantes e recombinantes são produzidos em proporções iguais. Esses quatro tipos de gametas se unem com o único tipo de gameta produzido pelo genitor homozigoto do cruzamento-teste para produzir quatro tipos de descendentes em proporções iguais (ver Figura 7.4 B). Os descendentes, com novas combinações de traços, formados a partir dos gametas

recombinantes são chamados de **descendentes recombinantes**.

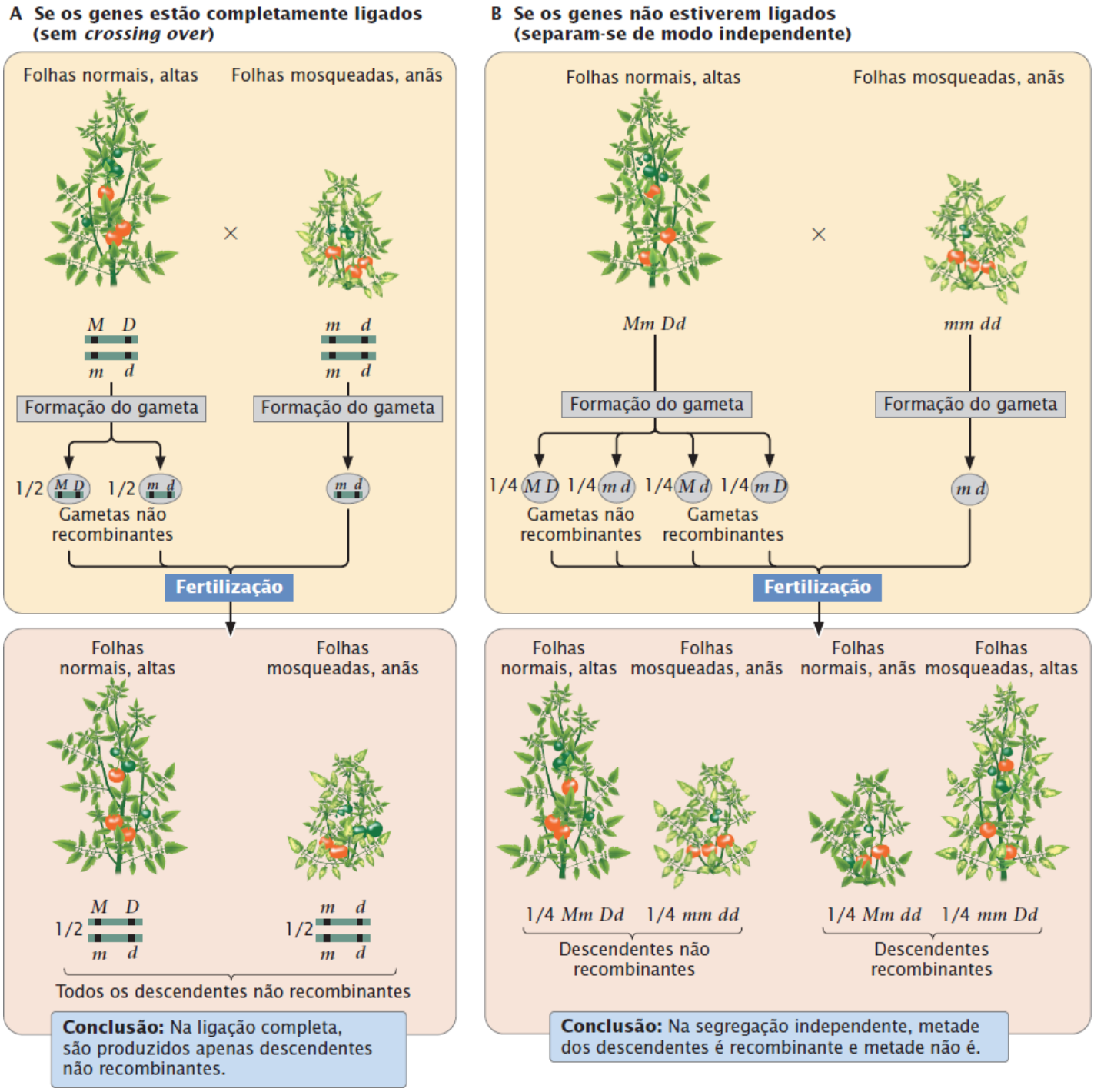


Figura 7.4 Um cruzamento-teste revela os efeitos da ligação. Os resultados de um cruzamento-teste para dois *loci* nos tomates que determinam o tipo de folha e a altura da planta.

Conceitos

Um cruzamento-teste no qual uma das plantas é heterozigota para dois genes completamente ligados gera dois tipos de descendentes, cada um apresentando uma das combinações originais dos traços encontrados na geração P. Em contrapartida, a segregação independente produz quatro tipos de descendentes em uma razão 1: 1: 1: 1 – dois tipos de descendentes recombinantes e dois tipos de descendentes não recombinantes em proporções iguais.

Crossing over com genes ligados

Em geral, ocorre *crossing over* entre genes que estão no mesmo cromossomo, produzindo novas combinações de traços. Os genes que exibem *crossing over* estão ligados de forma incompleta. Vamos ver como a ligação incompleta afeta os resultados de um cruzamento.

Teoria. O efeito do *crossing over* sobre a herança de dois genes ligados é mostrado na **Figura 7.5**. O *crossing over*, que ocorre na prófase I da meiose, é a troca de material genético entre cromátides não irmãs (ver Figuras 2.16 e 2.18). Após o *crossing over*, as duas cromátides que não participaram não são alteradas, de modo que os gametas que as recebem são não recombinantes. As outras duas, que participaram do *crossing over*, agora têm novas combinações de alelos e os gametas que as recebem são recombinantes. Para cada meiose na qual ocorre um *crossing over* único, serão produzidos dois gametas não recombinantes e dois recombinantes. Esse resultado é o mesmo que o produzido pela segregação independente (ver Figura 7.4 B); então, quando ocorre o *crossing over* entre dois *loci* em cada meiose, é impossível determinar se os genes estão no mesmo cromossomo e ocorre *crossing over*, ou se estão em diferentes cromossomos.

Para genes intimamente ligados, o *crossing over* não ocorre em cada meiose, sendo produzidos apenas gametas não recombinantes. Nas meioses com um *crossing over* único, metade dos gametas é recombinante e metade é não recombinante (porque um *crossing over* único afeta apenas duas de quatro cromátides). Como cada *crossing over* gera metade de gametas recombinantes e metade de gametas não recombinantes, a porcentagem total de gametas recombinantes é sempre a metade da porcentagem da meiose na qual ocorre *crossing over*. Mesmo se o *crossing over* entre dois genes ocorrer em cada meiose, apenas 50% dos gametas resultantes será recombinante. Assim, a frequência dos gametas recombinantes é sempre a metade da frequência do *crossing over* e a proporção máxima de gametas recombinantes é de 50%.

Conceitos

A ligação entre os genes faz com que eles sejam herdados juntos e reduz a combinação; o *crossing over* rompe as associações desses genes. Em um cruzamento-teste para dois genes ligados, cada *crossing over* produz dois gametas recombinantes e dois não recombinantes. A frequência dos gametas recombinantes é a metade da frequência do *crossing over*, e a frequência máxima de gametas recombinantes é 50%.

Checagem dos conceitos 1

Para *crossing overs* únicos, a frequência dos gametas recombinantes é metade da frequência do *crossing over* porque:

a. Um cruzamento-teste entre um homozigoto e um heterozigoto produz 1/2 de descendentes heterozigotos e 1/2 de descendentes homozigotos.
b. A frequência da recombinação é sempre 50%.
c. Cada *crossing over* ocorre entre apenas duas das quatro cromátides de um par homólogo.
d. Eles ocorrem em cerca de 50% das meioses.

Aplicação. Vamos aplicar o que aprendemos sobre ligação e recombinação para um cruzamento entre tomates com diferentes genes que codificam o tipo de folha e altura da planta. Considere agora que esses genes estão ligados e que ocorre *crossing over* entre eles. Suponha que um geneticista fez o cruzamento-teste descrito anteriormente:

$$\frac{M \quad D}{m \quad d} \times \frac{m \quad d}{m \quad d}$$

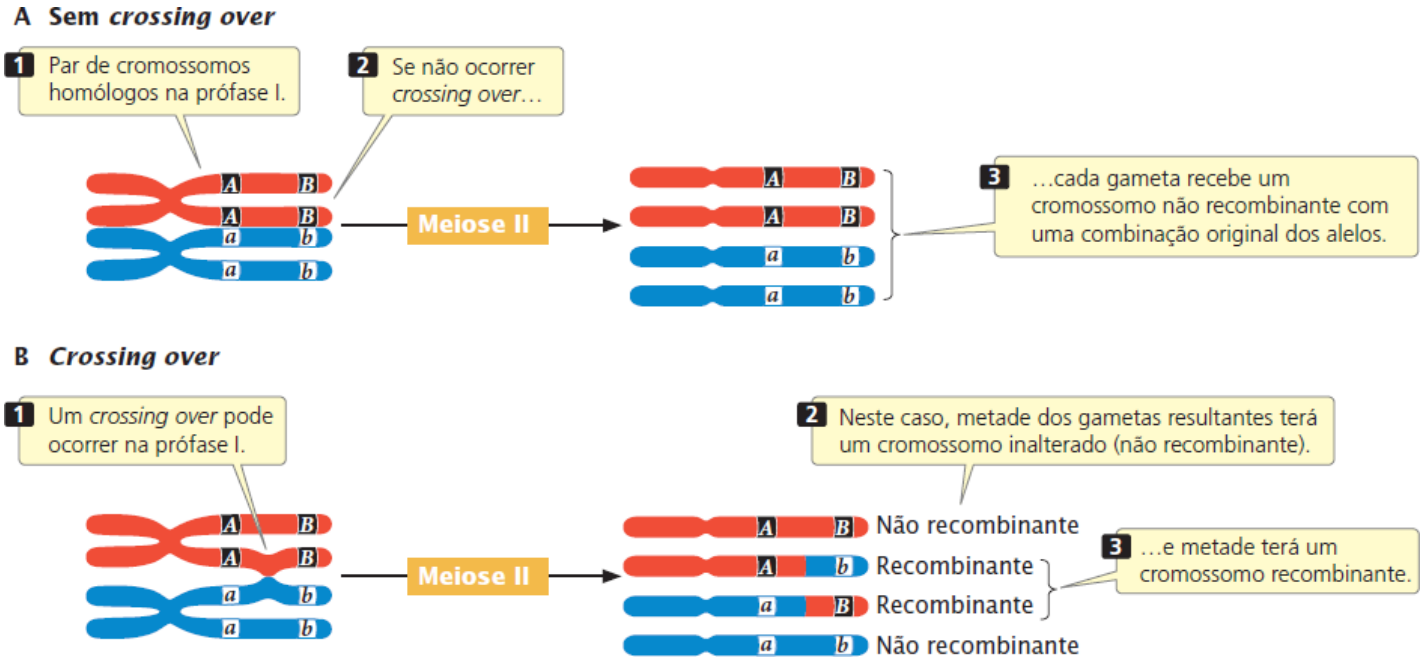


Figura 7.5 Um *crossing over* único produz metade dos gametas não recombinantes e metade dos gametas recombinantes.

Quando ocorre *crossing over* nos genes para o tipo de folha e altura, dois dos quatro gametas produzidos são recombinantes. Quando não há *crossing over*, todos os quatro gametas resultantes são não recombinantes. Como cada *crossing over* produz metade de gametas recombinante e metade não recombinante, a maioria dos gametas será não recombinante (**Figura 7.6 A**). Esses gametas então se unem aos produzidos pelo genitor homozigoto recessivo, que tem apenas os alelos recessivos, gerando a maioria de descendentes não recombinantes (**Figura 7.6 B**). Nesse cruzamento, vemos que 55 dos descendentes do cruzamento-teste têm folhas normais e são altos e 53 têm folhas mosqueadas e são anões. Essas plantas são os descendentes não recombinantes, com as combinações de traços originais encontrados nos genitores. Das 123 plantas descendentes, 15 têm novas combinações de traços que não eram observadas nos genitores: 8 têm folhas normais e são anãs e 7 têm folhas mosqueadas e são altas. Essas plantas são os descendentes recombinantes.

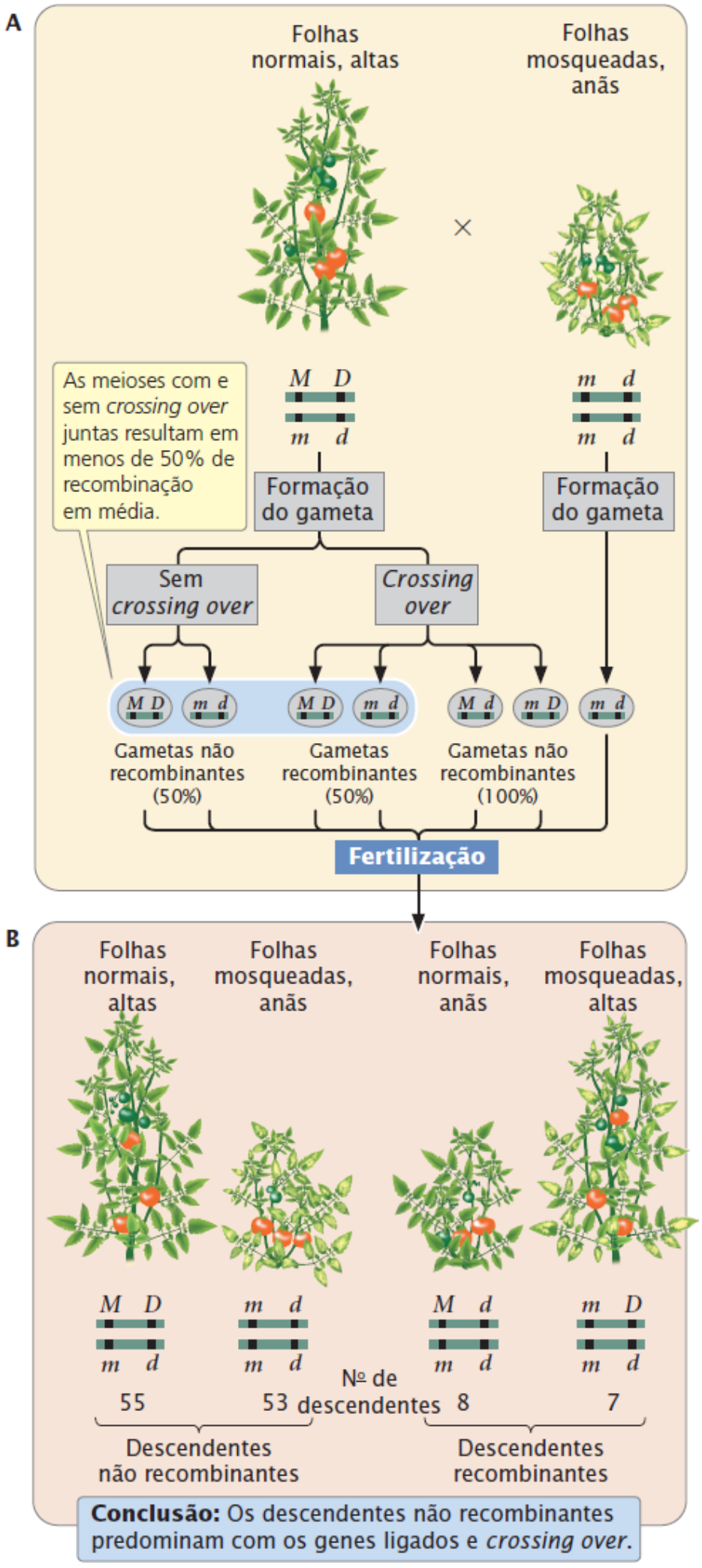

A
Folhas normais, altas
Folhas mosqueadas, anãs
×
As meioses com e sem *crossing over* juntas resultam em menos de 50% de recombinação em média.
M D
m d
m d
m d
Formação do gameta
Formação do gameta
Sem *crossing over*
Crossing over
M D
m d
M D
m d
M d
m D
m d
Gametas não recombinantes (50%)
Gametas recombinantes (50%)
Gametas não recombinantes (100%)
Fertilização
B
Folhas normais, altas
Folhas mosqueadas, anãs
Folhas normais, anãs
Folhas mosqueadas, altas
M D
m d
m d
m d
M d
m d
m D
m d
Nº de descendentes
55
53
8
7
Descendentes não recombinantes
Descendentes recombinantes
Conclusão: Os descendentes não recombinantes predominam com os genes ligados e *crossing over*.

Figura 7.6 O *crossing over* entre os genes ligados produz descendentes não recombinantes e recombinantes. Neste cruzamento-teste, os genes estão ligados e ocorre algum *crossing over*.

Os resultados de um cruzamento com o caso ilustrado na Figura 7.6 revelam vários itens. Um cruzamento-teste para dois genes que se separam de forma independente deve produzir uma razão fenotípica 1: 1: 1: 1 nos descendentes. Está claro que os descendentes desse cruzamento não exibem tal razão, então podemos suspeitar que os genes não se separam de forma independente. Quando os genes ligados sofrem *crossing over*, o resultado é a maioria de descendentes não recombinantes e poucos descendentes recombinantes. Este resultado é o que observamos entre os descendentes do cruzamento-teste ilustrado na Figura 7.6, então concluímos que os dois genes mostram o indício de ligação com algum *crossing over*.

Como calcular a frequência de recombinação

A porcentagem de descendentes recombinantes produzida por um cruzamento é chamada de **frequência de recombinação**, que é calculada pela fórmula a seguir:

$$\text{frequência de recombinação} = \frac{\text{nº de descendentes recombinantes}}{\text{nº total de descendentes}} \times 100\%$$

No cruzamento-teste mostrado na Figura 7.6, quinze descendentes exibem novas combinações de traços, então a frequência de recombinação é:

$$\frac{8+7}{55+53+8+7} \times 100\% = \frac{15}{123} \times 100\% = 12{,}2\%$$

Assim, 12,2% dos descendentes exibem novas combinações de traços resultantes do *crossing over*. A frequência de recombinação também pode ser expressa como decimal (0,122). **Resolva o Problema 15**

Acoplamento e repulsão

Nos cruzamentos para genes ligados, a disposição dos alelos nos cromossomos homólogos é crítica para determinar o desfecho do cruzamento. Por exemplo, considere a herança de dois genes na mosca-varejeira australiana, *Lucilia cuprina*. Na espécie, um *locus* determina a cor do tórax: um tórax púrpura (*p*) é recessivo para o tórax verde normal (p^+). Um segundo *locus* determina a cor do pupário: um pupário preto (*b*) é recessivo para o pupário marrom normal (b^+). Os *loci* para a cor do tórax e do pupário estão próximos no cromossomo. Suponha que realizamos um cruzamento-teste de uma mosca que seja heterozigota em ambos os *loci* com uma mosca que seja homozigota recessiva para ambos. Como esses genes estão ligados, existem duas possíveis disposições nos cromossomos das moscas descendentes heterozigotas. Os alelos dominantes para o tórax verde (p^+) e o pupário marrom (b^+) podem residir em um cromossomo do par homólogo e os alelos recessivos para o tórax púrpura (*p*) e pupário preto (*b*) podem residir no outro cromossomo homólogo:

$$\frac{p^+ \quad b^+}{p \quad b}$$

Essa disposição, na qual os alelos do tipo selvagem são encontrados em um cromossomo e os mutantes no outro, é chamada de **acoplamento** ou **configuração *cis***. Por outro lado, um cromossomo pode carrear os alelos para o tórax verde (p^+) e o pupário preto (*b*) e outro os alelos para tórax púrpura (*p*) e pupário marrom (b^+):

$$\frac{p^+ \quad b}{p \quad b^+}$$

Essa disposição, na qual cada cromossomo tem um alelo do tipo selvagem e um alelo mutante, é chamada de **repulsão** ou **configuração *trans***. O fato de os alelos no genitor heterozigoto estarem em acoplamento ou repulsão determina quais fenótipos serão mais comuns entre os descendentes de um cruzamento-teste.

Quando os alelos estão na configuração de acoplamento, a maioria dos tipos de descendentes são os que apresentam tórax verde e pupário marrom e os que apresentam tórax púrpura e pupário preto (**Figura 7.7 A**). Quando os alelos do genitor heterozigoto estão na configuração de repulsão, a maioria dos tipos de descendentes são os que apresentam tórax verde e pupário preto e os que apresentam tórax púrpura e pupário marrom (**Figura 7.7 B**). Observe que os genótipos dos genitores na Figura 7.7 A e B são os mesmos ($p^+ p\ b^+ b \times pp\ bb$) e que a drástica diferença nas razões fenotípicas dos descendentes nos dois cruzamentos é um resultado da configuração – acoplamento ou repulsão dos cromossomos. O conhecimento da disposição dos alelos nos cromossomos é essencial para prever com exatidão o desfecho dos cruzamentos nos quais os genes estão ligados.

Conceitos

Em um cruzamento, a disposição dos alelos ligados nos cromossomos é crítica para determinar o desfecho. Quando dois alelos do tipo selvagem estão em um cromossomo homólogo e dois alelos mutantes estão em outro, eles estão na configuração de acoplamento; quando cada cromossomo tem um alelo do tipo selvagem e um mutante, eles estão em repulsão.

Checagem dos conceitos 2

O seguinte cruzamento-teste produz estes descendentes: *Aa Bb* × *aa bb* → 10 *Aa Bb*, 40 *Aa bb*, 40 *aa Bb*, 10 *aa bb*. Os alelos *A* e *B* no genitor *Aa Bb* estão em acoplamento ou repulsão?

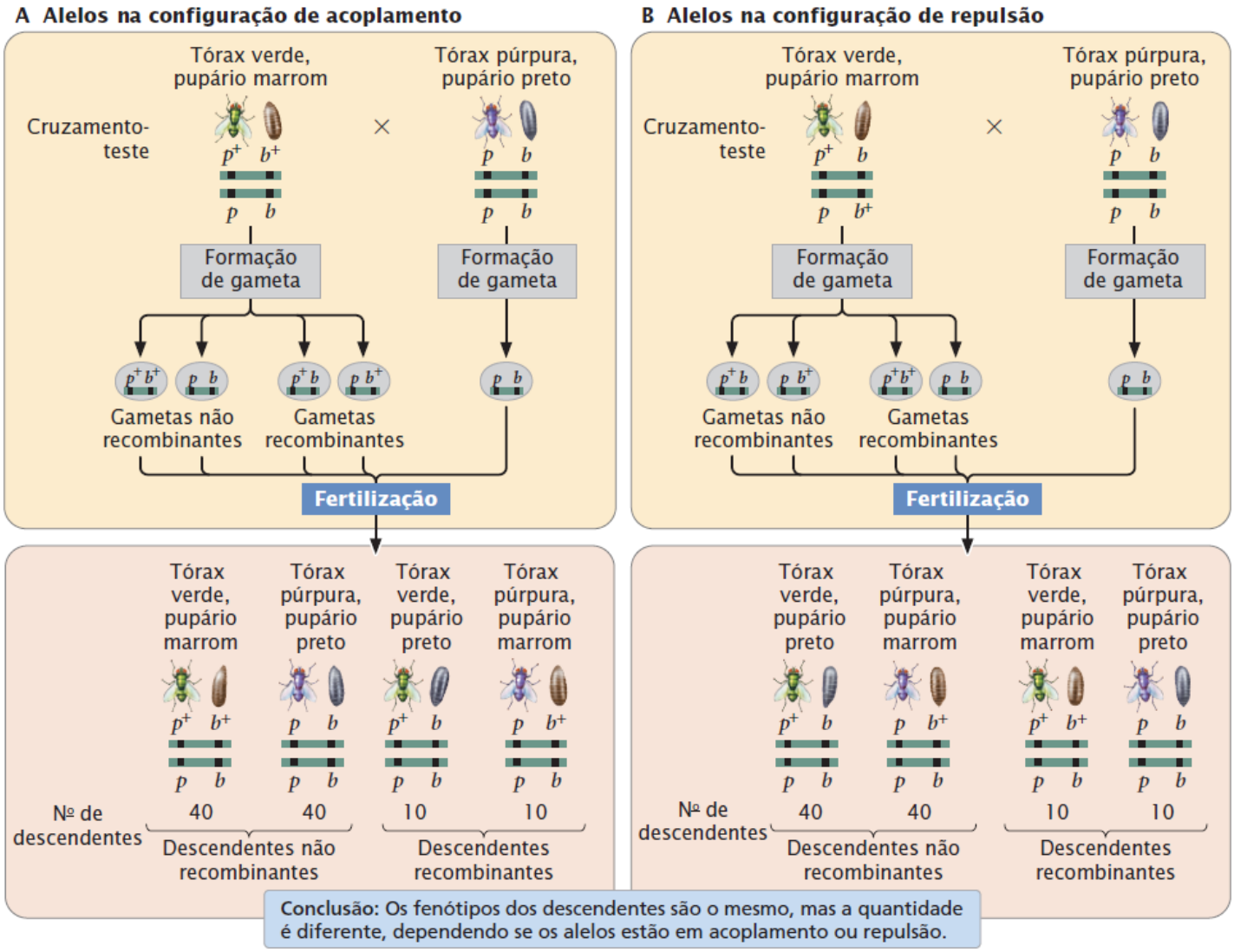


Figura 7.7 A disposição (acoplamento ou repulsão) dos genes ligados em um cromossomo afeta os resultados de um cruzamento-teste. Os *loci* ligados na mosca-varejeira australiana, *Lucilia cuprina*, determinam a cor do tórax e do pupário.

Conceitos conectantes

Relacionamento de segregação independente, ligação e *crossing over*

Agora vamos considerar três situações envolvendo os genes em diferentes *loci*. Primeiro, os genes podem ser localizados em diferentes cromossomos; neste caso, eles exibem a segregação independente e se combinam de forma aleatória quando os gametas são formados. Um heterozigoto em dois *loci* (*Aa Bb*) produz quatro tipos de gametas (*A B*, *a b*, *A b* e *a B*) em proporções iguais: dois tipos de não recombinantes e dois tipos de recombinantes. Em um cruzamento-teste, esses gametas resultarão em quatro tipos de descendentes em proporções iguais (**Quadro 7.1**).

Segundo, os genes podem estar completamente ligados – o que significa que eles estão no mesmo cromossomo e tão próximos que o *crossing over* entre eles é raro. Neste caso, os genes não se recombinam. Um heterozigoto para dois genes intimamente ligados na configuração de acoplamento

$\frac{A \quad B}{a \quad b}$

produz apenas gametas não recombinantes contendo os alelos *A B* ou *a b*, não se separando em novas combinações como *A b* e *a B*. Em um cruzamento-teste, os genes completamente ligados produzem apenas dois tipos de descendentes, ambos não recombinantes, em proporções iguais (ver Quadro 7.1).

A terceira situação, a ligação incompleta, é intermediária entre os dois extremos de segregação independente e ligação completa. Aqui, os genes estão fisicamente ligados no mesmo cromossomo, o que evita a segregação independente. Entretanto, *crossing overs* ocasionais rompem as ligações e permitem que os genes se recombinem. Com a ligação incompleta, um heterozigoto em dois *loci* produz quatro tipos de gametas – dois tipos de recombinantes e dois tipos de não recombinantes – mas os não recombinantes são produzidos com maior frequência que os recombinantes porque o *crossing over* não ocorre em todas as meioses. No cruzamento-teste, esses gametas resultam em quatro tipos de descendentes, com os não recombinantes mais frequentes que os recombinantes (ver Quadro 7.1).

No início do capítulo, o termo recombinante foi definido como a seleção dos alelos em novas combinações. Agora, vamos considerar dois tipos de recombinação com diferentes mecanismos para produzir essas novas combinações de alelos. A recombinação intercromossômica ocorre entre os genes localizados em *diferentes* cromossomos. Ela surge a partir da segregação independente – a separação aleatória dos cromossomos na anáfase I da meiose – e é o tipo que Mendel descobriu enquanto estudava os cruzamentos di-híbridos. Um segundo tipo de recombinação, a recombinação intracromossômica, ocorre entre genes localizados no *mesmo* cromossomo. Ela surge a partir do *crossing over* – a troca de material genético na prófase I da meiose. Ambos os tipos de recombinação produzem novas combinações de alelos nos gametas de modo que eles não podem ser distinguidos ao serem examinados. Contudo, eles podem ser distinguidos pelas *frequências* dos tipos de gametas: a recombinação intercromossômica produz 50% de gametas não recombinantes e 50% de recombinantes, enquanto a recombinação intracromossômica produz mais de 50% de gametas não recombinantes e menos de 50% de recombinantes. Entretanto, quando os genes estão muito afastados no mesmo cromossomo, o *crossing over* ocorre em cada divisão meiótica, gerando 50% de gametas recombinantes e 50% de não recombinantes. Esse resultado é o mesmo da segregação independente dos genes localizados em diferentes cromossomos (recombinação intercromossômica). Assim, a recombinação intracromossômica dos genes que estão afastados no mesmo cromossomo e a recombinação intercromossômica são indistinguíveis do ponto de vista do fenótipo.

Quadro 7.1 Resultados de um cruzamento-teste (*Aa Bb* × *aa bb*) com ligação completa, segregação independente e ligação com algum *crossing over*.

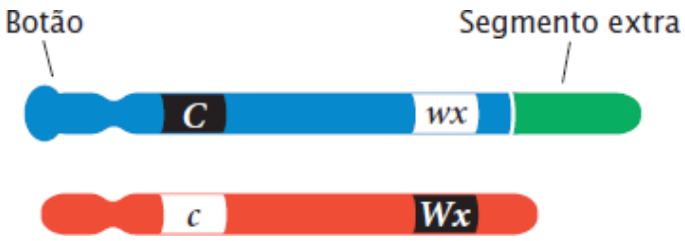
Situação	Descendentes do cruzamento-teste	
Segregação independente	*Aa Bb* (não recombinante)	25%
	aa bb (não recombinante)	25%
	Aa bb (recombinante)	25%
	aa Bb (recombinante)	25%
Ligação completa (genes em acoplamento)	*Aa Bb* (não recombinante)	50%
	aa bb (não recombinante)	50%
Ligação com algum *crossing over* (genes em acoplamento)	*Aa Bb* (não recombinante) *aa bb* (não recombinante)	Mais que 50%
	Aa bb (recombinante) *aa Bb* (recombinante)	Menos que 50%

Evidência para a base física da recombinação

A teoria cromossômica da herança de Walter Sutton, segundo a qual os genes estão fisicamente localizados nos cromossomos, foi apoiada pela descoberta de Nettie Stevens e Edmund Wilson de que o sexo estava associado a um cromossomo específico nos insetos e pela demonstração de Calvin Bridges de que a não disjunção dos cromossomos X estava relacionada com a herança da cor dos olhos na *Drosophila* (ver *Capítulo 4*). Surgiram mais evidências para a teoria cromossômica da herança em 1931, quando Harriet Creighton e Barbara McClintock obtiveram mais indício de que a recombinação intracromossômica era o resultado da troca física entre os cromossomos. Creighton e McClintock descobriram uma cepa de milho que tinha um cromossomo 9 anormal, contendo um

botão de coloração intensa em uma extremidade e um pequeno pedaço de outro cromossomo preso a outra. Esse cromossomo anômalo possibilitou que eles distinguissem visualmente os dois membros de um par homólogo.

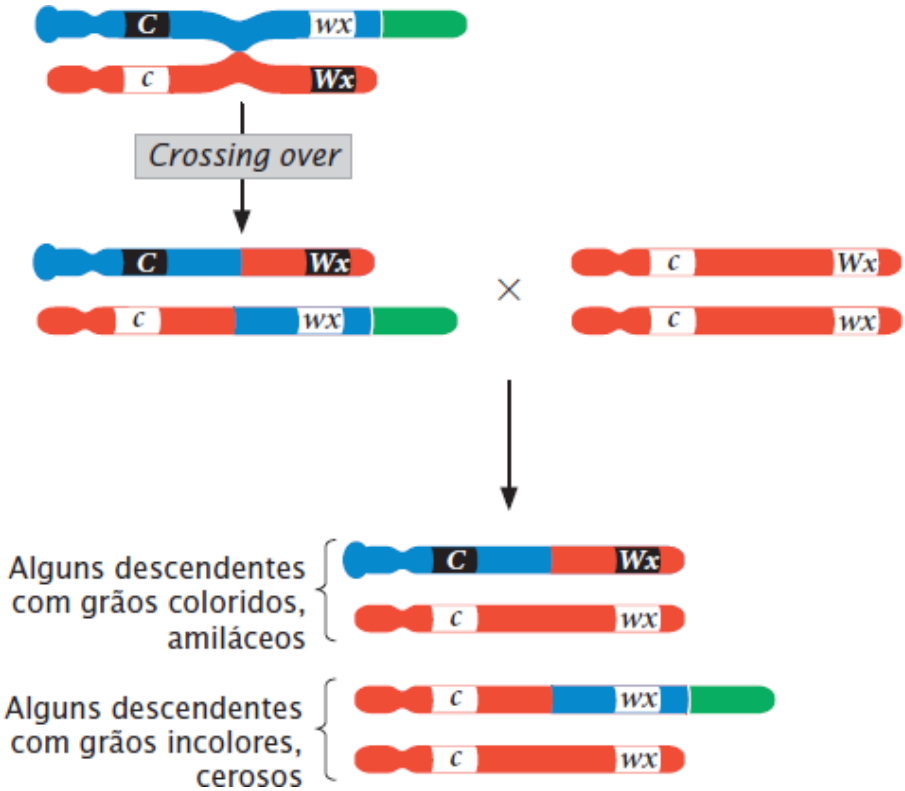
Creighton e McClintock estudaram a herança de dois traços no milho determinados pelos genes no cromossomo 9. Em um *locus*, um alelo dominante (*C*) produzia grãos coloridos, enquanto um alelo recessivo (*c*) produzia grãos incolores. Em um segundo *locus*, um alelo dominante (*Wx*) produzia grãos amiláceos, enquanto um alelo recessivo (*wx*) produzia grãos cerosos. Eles obtiveram uma planta heterozigota em ambos os *loci* em repulsão, com os alelos para grão colorido e ceroso no cromossomo anômalo e os alelos para grão incolor e amiláceo no cromossomo normal:



Elas, então, cruzaram a planta heterozigota com uma homozigota para grão incolor e heterozigota para grão ceroso (com ambos os cromossomos normais):

$$\frac{C \quad wx}{c \quad Wx} \times \frac{c \quad Wx}{c \quad wx}$$

Este cruzamento produzirá diferentes combinações de traços nos descendentes, mas a única forma para que os descendentes de grão incolor e ceroso possam surgir é por *crossing over* duplo no genitor heterozigoto:



Observe que, se o *crossing over* possibilita a troca física entre os cromossomos, então os descendentes de grãos incolores e cerosos, resultantes da recombinação, devem ter um cromossomo com um pedaço extra, mas sem um botão. Além disso, alguns dos descendentes com grão colorido e amiláceo devem ter um botão, mas não o pedaço extra. Este desfecho é exatamente o que Creighton e McClintock observaram, confirmando a teoria cromossômica da herança. Curt Stern forneceu uma demonstração semelhante ao usar os marcadores cromossômicos na *Drosophila* ao mesmo tempo. Vamos examinar a base molecular da recombinação com mais detalhes no Capítulo 12.

Como prever os desfechos dos cruzamentos com genes ligados

Conhecer a disposição dos alelos em um cromossomo permite que possamos prever os tipos de descendentes que resultarão de um cruzamento que transmite genes ligados e determinar quais desses tipos serão os mais numerosos. Determinar as *proporções* dos tipos de descendentes requer uma informação adicional – a frequência da recombinação. Ela fornece informações sobre com que

frequência os alelos nos gametas surgem em novas combinações, possibilitando que possamos prever as proporções de fenótipos dos descendentes que surgirão a partir de um cruzamento específico com genes ligados.

Nos pepinos, a fruta lisa (*t*) é recessiva para a fruta rugosa (*T*) e a fruta brilhosa (*d*) é recessiva para a fruta fosca (*D*). Os geneticistas determinaram que esses dois genes exibem uma frequência de recombinação de 16%. Suponha que cruzemos uma planta homozigota para fruta rugosa e fosca com uma planta homozigota para fruta lisa e brilhosa e, então, façamos um cruzamento-teste com a F_1:

$$\frac{T \quad D}{t \quad d} \times \frac{t \quad d}{t \quad d}$$

Quais tipos e proporções dos descendentes são esperados a partir do cruzamento-teste?

Quatro tipos de gametas serão produzidos pelo genitor heterozigoto, como demonstrado na **Figura 7.8**: dois não recombinantes (*T D* e *t d*) e dois tipos recombinantes (*T d* e *t D*). A frequência de recombinação informa que 16% dos gametas produzidos pelo genitor heterozigoto serão recombinantes. Como existem dois tipos de gametas recombinantes, cada um deve surgir em uma frequência de 16%/2 = 8%. Essa frequência também pode ser expressa como uma probabilidade (0,08). Todos os outros gametas serão não recombinantes, então eles devem surgir em uma frequência de 100% – 16% = 84%. Como existem dois tipos de gametas não recombinantes, cada um deve surgir em uma frequência de 84%/2 = 42% (ou 0,42). O outro genitor no cruzamento-teste é homozigoto e, portanto, produz apenas um único tipo de gameta (*t d*) com uma frequência de 100% (ou 1,00).

Quatro tipos de descendentes surgem a partir do cruzamento-teste (ver Figura 7.8). A proporção esperada de cada tipo pode ser determinada ao usar a regra de multiplicação (ver *Capítulo 3*), ao multiplicar a probabilidade de cada gameta. Os descendentes do cruzamento-teste com fruta rugosa e fosca

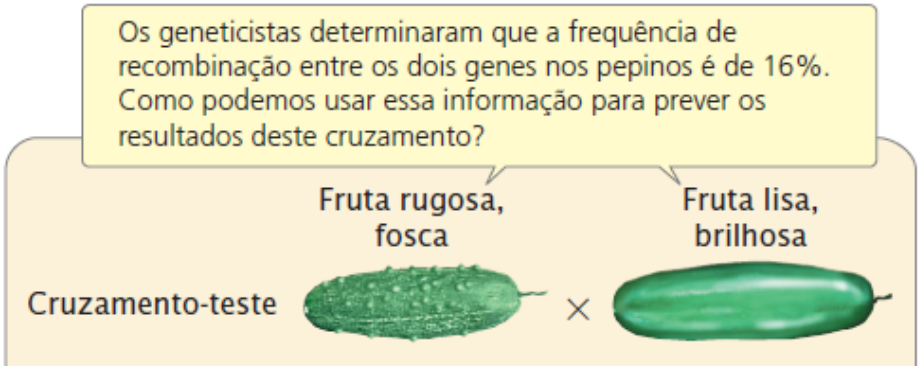
$$\frac{T \quad D}{t \quad d}$$

aparecem com uma frequência de 0,42 (a probabilidade de herdar um gameta com cromossomo *T D* a partir do genitor heterozigoto) × 1,00 (a probabilidade de herdar um gameta com cromossomo *t d*) do genitor recessivo) = 0,42. As proporções de outros tipos de descendentes F_2 podem ser calculadas de forma semelhante (ver Figura 7.8). Esse método pode ser usado para prever o desfecho de qualquer resultado com genes ligados para os quais a frequência de recombinação é conhecida.

Análise de segregação independente

Em alguns cruzamentos, os genes estão claramente ligados porque existem mais descendentes não recombinantes que recombinantes. Em outros, a diferença entre a segregação independente e a ligação não está tão clara. Por exemplo, suponha que se faça um cruzamento-teste para dois pares de genes, como *Aa Bb* × *aa bb*, e observemos os seguintes descendentes: 54 *Aa Bb*, 56 *aa bb*, 42 *Aa bb* e 48 *aa Bb*. Essa é a razão 1: 1: 1: 1, que esperaríamos se *A* e *B* se separassem de forma independente? Não exatamente, mas bem perto. Talvez, esses genes se separem de forma independente e a mudança produziu sutis desvios entre os números e a razão 1: 1: 1: 1 esperados. Por outro lado, os genes podem estar ligados, com considerável *crossing over* ocorrendo entre eles, e o número de não recombinantes é um pouco maior que o de recombinantes. Como distinguimos entre o papel do acaso e o papel da ligação para produzir desvios dos resultados esperados com a segregação independente?

Encontramos problema semelhante nos cruzamentos nos quais os genes não estavam ligados – o problema de distinguir entre os desvios aleatórios e os causados por outros fatores. Tratamos esse caso (no *Capítulo 3*) com o teste qui-quadrado de adequação de ajuste, que nos ajuda a avaliar a probabilidade de que o acaso sozinho seja responsável pelos desvios entre os números de descendentes que observamos e o número que esperamos ao aplicar os princípios da hereditariedade. Aqui, estamos interessados em uma questão diferente: a herança dos alelos em um *locus* é independente da herança dos alelos em um segundo *locus*? Se a resposta a essa pergunta for sim, então os genes se separam de forma independente; se a resposta for não, então é provável que os genes estejam ligados.



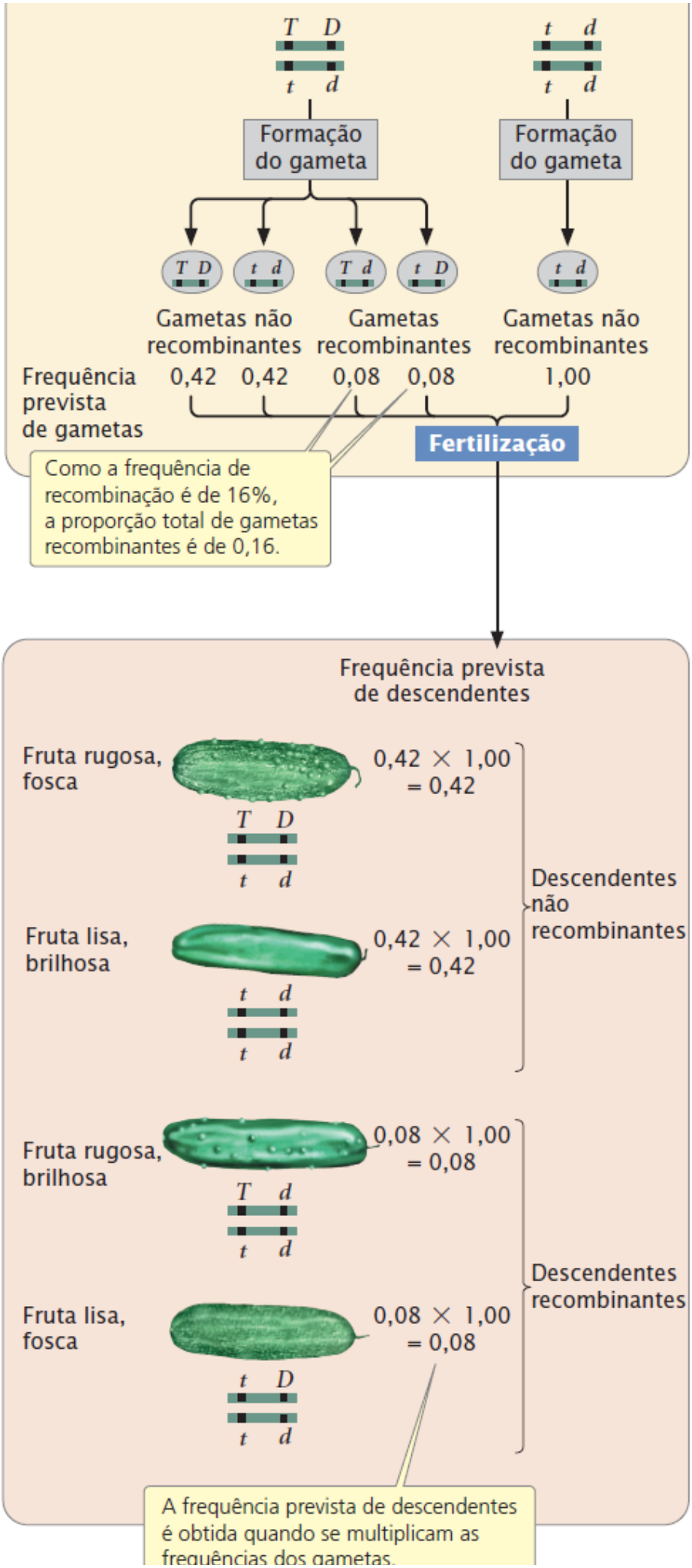

T D
t d
t d
t d
Formação do gameta
Formação do gameta
T D
t d
T d
t D
t d
Gametas não recombinantes
Gametas recombinantes
Gametas não recombinantes
Frequência prevista de gametas
0,42
0,42
0,08
0,08
1,00
Fertilização
Como a frequência de recombinação é de 16%, a proporção total de gametas recombinantes é de 0,16.
Frequência prevista de descendentes
Fruta rugosa, fosca
T D
t d
0,42 × 1,00 = 0,42
Fruta lisa, brilhosa
t d
t d
0,42 × 1,00 = 0,42
Descendentes não recombinantes
Fruta rugosa, brilhosa
T d
t d
0,08 × 1,00 = 0,08
Fruta lisa, fosca
t D
t d
0,08 × 1,00 = 0,08
Descendentes recombinantes
A frequência prevista de descendentes é obtida quando se multiplicam as frequências dos gametas.

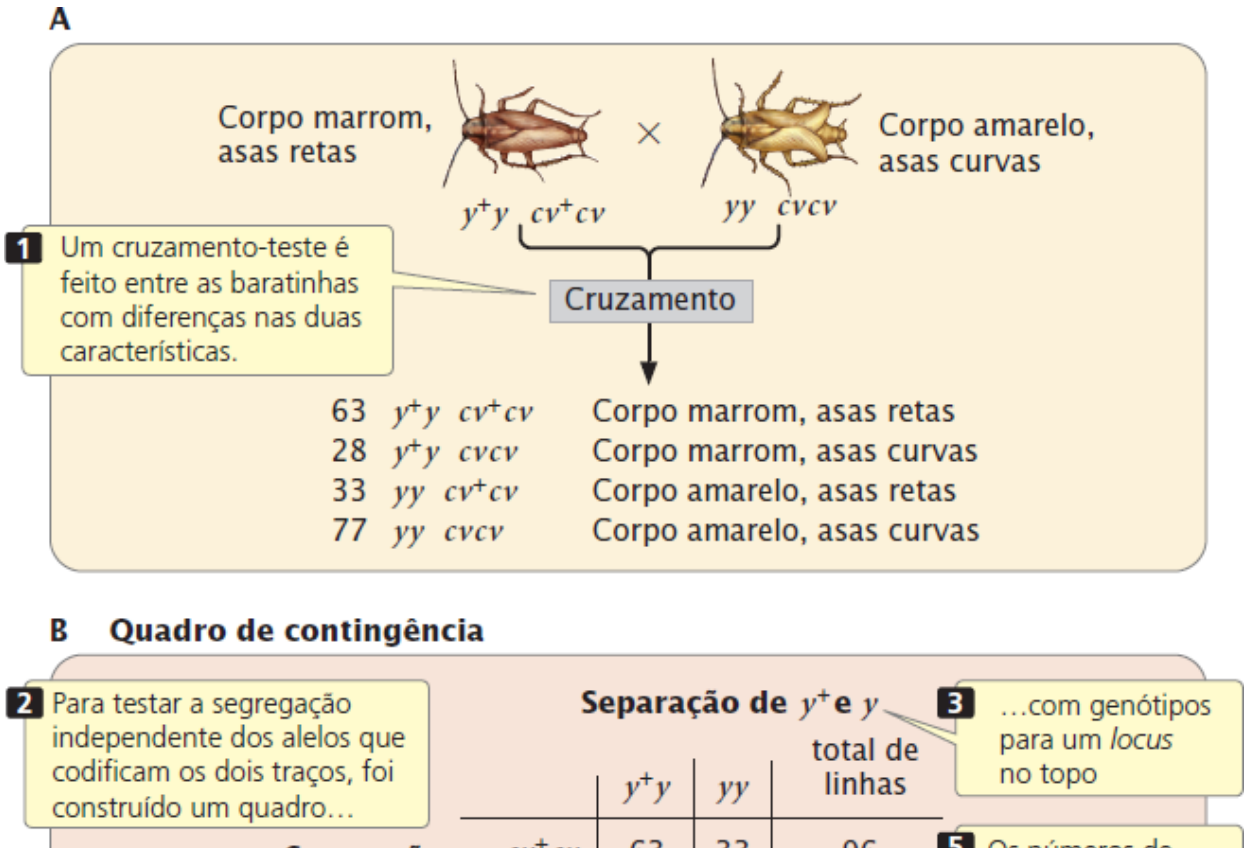
Figura 7.8 A frequência de recombinação permite uma previsão das proporções de descendentes esperadas para um cruzamento que transmite genes ligados.

Uma forma possível para testar a segregação independente é calcular a probabilidade esperada de cada tipo de descendente, considerando a segregação independente, e, então, usar o teste qui-quadrado de adequação de ajuste para avaliar se os números observados se desviam significativamente dos esperados. Com a segregação independente, esperamos 1/4 de cada fenótipo: 1/4 *Aa Bb*, 1/4 *aa bb*, 1/4 *Aa bb* e 1/4 *aa Bb*. Essa probabilidade se baseia na regra de multiplicação de probabilidade (ver *Capítulo 3*). Por exemplo, se a probabilidade de *Aa* é 1/2 e a de *Bb* é 1/2, então a probabilidade de *Aa Bb* é 1/2 × 1/2 = 1/4. Neste cálculo, fazemos suas suposições: (1) a probabilidade de cada genótipo de único *locus* é 1/2 e (2) os genótipos em dois *loci* são herdados de forma independente (1/2 × 1/2 = 1/4).

Um problema com esta abordagem é que um valor significativo de qui-quadrado pode surgir a partir da violação de uma das suposições. Se os genes estão ligados, então a herança dos genótipos em dois *loci* não é independente (hipótese 2) e teremos um desvio significativo entre os valores observados e esperados. Mas também podemos ter um desvio significativo se a probabilidade do genótipo de único *locus* não for 1/2 (hipótese 1), mesmo quando os genótipos se separam de forma independente. Podemos obter um desvio significativo, por exemplo, se os indivíduos com um genótipo tiverem menor probabilidade de sobreviver ou a penetrância de um genótipo não for 100%. Podemos testar ambas as hipóteses ao realizar vários testes qui-quadrado, primeiro testando a herança dos genótipos em cada *locus* separadamente (hipótese 1) e, então, testar para segregação independente (hipótese 2). Entretanto, um método mais rápido é testar a independência nos genótipos com um *teste qui-quadrado de independência*.

Teste qui-quadrado de independência. O teste qui-quadrado de independência permite que avaliemos se a separação dos alelos em um *locus* é independente da separação dos alelos em outro *locus*, sem fazer nenhuma suposição sobre a probabilidade dos genótipos de único *locus*. Para ilustrar esta análise, examinaremos os resultados a partir de um cruzamento entre baratinhas, nas quais o corpo amarelo (*y*) é recessivo para o corpo marrom (y^+) e as asas curvas (*cv*) são recessivas para as asas retas (cv^+). Um cruzamento-teste ($y^+y\ cv^+cv \times yy\ cvcv$) produziu os descendentes apresentados na **Figura 7.9 A**. Se a separação dos alelos em cada *locus* é independente, então a proporção de descendentes com os genótipos y^+y e *yy* deve ser a mesma para as baratas-alemães com genótipo cv^+cv e *cvcv*. O inverso também é verdadeiro: as proporções de descendentes com os genótipos cv^+cv e *cvcv* devem ser as mesmas para as baratinhas com genótipo y^+y e *yy*.

Para determinar se as proporções de descendentes com genótipos em dois *loci* são independentes, primeiro construímos um quadro dos números observados dos descendentes, como um quadrado de Punnett, exceto que vamos colocar os genótipos que surgem a partir da separação dos alelos em um *locus* ao longo do topo e os genótipos que surgem a partir da separação dos alelos em outro *locus* na lateral (**Figura 7.9 B**). Depois, computamos o total para cada linha, o total para cada coluna e o total final (a soma de todos os totais das linhas ou a soma de todos os totais das colunas, que devem ser o mesmo). Esses totais serão usados para computar os valores esperados para o teste qui-quadrado de independência.



Separação de cv^+ e cv			
$cv^+ cv$	63	33	96
$cv\ cv$	28	77	105
Total de colunas	91	110	201
			Total final

4 …e os genótipos para o outro *locus* no lado esquerdo.

5 Os números de cada genótipo são colocados nas células do quadro e os totais de linha, totais de coluna e total final são computados.

C

Genótipo	Número observado	Número esperado $\left(\frac{\text{Total de linhas} \times \text{Total de colunas}}{\text{Total final}}\right)$
$y^+y\ cv^+cv$	63	$\frac{96 \times 91}{201} = 43{,}46$
$y^+y\ cvcv$	28	$\frac{105 \times 91}{201} = 47{,}54$
$yy\ cv^+cv$	33	$\frac{96 \times 110}{201} = 52{,}46$
$yy\ cvcv$	77	$\frac{105 \times 110}{201} = 57{,}46$

6 Os números esperados de descendentes, assumindo a separação independente, são calculados.

D

$$\chi^2 = \Sigma \frac{(\text{Observados} - \text{Esperados})}{\text{Esperados}}$$

$$= \frac{(63 - 43{,}46)^2}{43{,}46} + \frac{(28 - 47{,}54)^2}{47{,}54} + \frac{(33 - 52{,}54)^2}{52{,}54} + \frac{(77 - 57{,}46)^2}{57{,}46}$$

$$= 8{,}79 + 8{,}03 + 7{,}27 + 6{,}64$$

$$= 30{,}73$$

7 É calculado um valor de qui-quadrado.

E

gl = (Número de linhas − 1) × (Número de colunas − 1)

$gl = (2 - 1) \times (2 - 1) = 1 \times 1 = 1$

$P < 0{,}005$

8 A probabilidade é menos de 0,005, indicando que é provável que a *diferença* entre os números dos descendentes observados e esperados não seja aleatória.

Conclusão: Os genes para cor do corpo e tipo de asa não se separam de maneira independente e devem estar ligados.

Figura 7.9 Um teste qui-quadrado de independência pode ser usado para determinar se os genes em dois *loci* estão se separando de maneira independente.

Nosso próximo passo é computar os valores esperados para cada combinação dos genótipos (cada célula no quadro) com a hipótese de que a separação dos alelos no *locus y* é independente da separação dos alelos no *locus cv*. Se a separação dos alelos em cada *locus* for independente, o número esperado em cada célula pode ser computado com a seguinte fórmula:

$$\text{Número esperado} = \frac{\text{Total de linhas} \times \text{Total de colunas}}{\text{Total final}}$$

Para a célula do quadro correspondente ao genótipo y^+y^+ *cvcv* (a célula do canto à esquerda no quadro na Figura 7.9 B), o número esperado é:

$$\frac{96\ (\text{Total de linhas}) \times 91\ (\text{Total de colunas})}{201\ (\text{Total final})} = \frac{8.736}{201} = 43{,}46$$

Com o uso deste método, os números esperados para cada célula estão indicados na **Figura 7.9 C**.

Agora calculamos um valor de qui-quadrado ao usar a mesma fórmula que usamos para o teste qui-quadrado de adequação de ajuste, no Capítulo 3:

$$x^2 = \sum \frac{(\text{Observados} - \text{Esperados})^2}{\text{Esperados}}$$

Observe que Σ significa o "somatório" e que estamos adicionando o valor (observado – esperado)^2/valor esperado para cada tipo de descendente. Com os números observados e esperados de baratas-alemães a partir do cruzamento-teste, o valor de qui-quadrado calculado é 30,73 (**Figura 7.9 D**).

Para determinar a probabilidade associada a esse valor de qui-quadrado, precisamos ter os graus de liberdade. Lembre-se do Capítulo 3 e de que os graus de liberdade são o número de formas nas quais as classes observadas são livres para variar dos valores esperados. Em geral, para o teste qui-quadrado de independência, os graus de liberdade são iguais ao número de linhas no quadro menos 1, multiplicados pelo número de colunas no quadro menos 1 (**Figura 7.9 E**) ou:

$$gl = (\text{Número de linhas} - 1) \times (\text{Número de colunas} - 1)$$

Em nosso exemplo, existem duas linhas e colunas, então os graus de liberdade são:

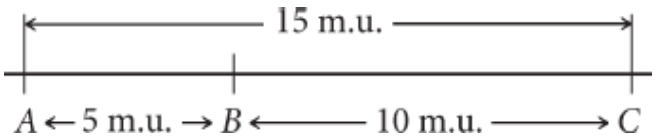
$$gl = (2 - 1) \times (2 - 1) = 1 \times 1 = 1$$

Portanto, nosso valor de qui-quadrado calculado é 30,73, com 1 grau de liberdade. Podemos usar o Quadro 3.7 para encontrar a probabilidade associada. Consultando o Quadro 3.7, podemos descobrir que nosso valor de qui-quadrado calculado é maior que o maior valor de qui-quadrado fornecido para 1 grau de liberdade, que tem uma probabilidade de 0,005. Assim, nosso valor de qui-quadrado calculado tem uma probabilidade menor que 0,005, muito baixa, indicando que os genótipos não estão nas proporções que esperaríamos se a segregação independente estivesse ocorrendo. Nossa conclusão, então, é de que os genes não estão se separando de forma independente e devem estar ligados. Como é o caso para o teste qui-quadrado de adequação de ajuste, em geral os geneticistas consideram que qualquer valor de qui-quadrado para o teste de independência com uma probabilidade menor que 0,05 é muito diferente dos valores esperados e, portanto, é evidência de que os genes não se separam de forma independente. **Resolva o Problema 16**

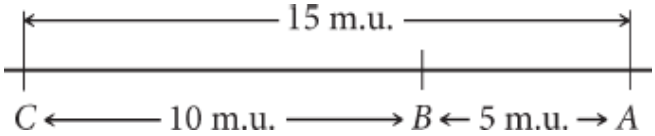
Mapeamento de genes com frequências de recombinação

Thomas Hunt Morgan e seus estudantes desenvolveram a ideia de que as distâncias físicas entre os genes em um cromossomo estão relacionadas com as taxas de recombinação. Eles criaram uma hipótese de que os eventos de *crossing over* ocorrem mais ou menos de forma aleatória para cima e para baixo no cromossomo e que é mais provável que dois genes distantes um dos outros sofram um *crossing over* que dois genes que estão próximos. Eles propuseram que as frequências de recombinação poderiam ser uma forma conveniente para determinar a ordem dos genes em um cromossomo e daria uma estimativa das distâncias relativas entre eles. Os mapas cromossômicos calculados usando o fenômeno genético de recombinação são chamados de **mapas genéticos**. Por outro lado, os mapas cromossômicos calculados ao usar as distâncias físicas no cromossomo (frequentemente expressos como números de pares de bases) são chamados de **mapas físicos**.

As distâncias nos mapas genéticos são medidas em **unidades de mapa** (abreviadas como m.u.), sendo uma unidade de mapa igual a 1% de recombinação. As unidades de mapa também são chamadas de **centiMorgans** (cM) em homenagem a Thomas Hunt Morgan. As distâncias genéticas medidas com as taxas de recombinação são aproximadamente aditivas: se a distância do gene *A* para o gene *B* é 5 m.u., a distância do gene *B* para o gene *C* é 10 m.u. e a distância do gene *A* para o gene *C* é 15 m.u., então o gene *B* deve estar localizado entre os genes *A* e *C*. Com base nas distâncias do mapa fornecidas, podemos desenhar um mapa genético simples para os genes *A*, *B* e *C*, como é demonstrado a seguir:



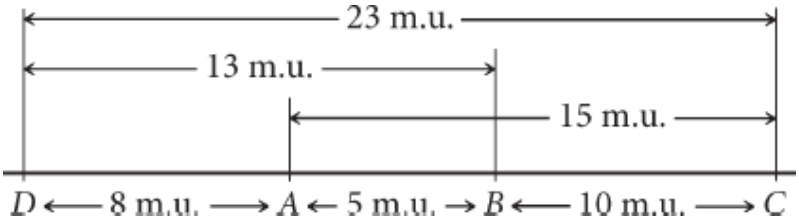
Poderíamos de forma plausível desenhar este mapa com *C* à esquerda e *A* à direita:



Ambos os mapas estão corretos e são equivalentes porque, com as informações sobre as posições relativas de apenas três genes, o máximo que podemos determinar é qual gene está no meio. Se tivermos as distâncias para um gene adicional, então podemos posicionar *A* e *C* em relação a esse gene. Um gene *D* adicional, examinado por meios de cruzamentos genéticos, pode gerar as seguintes frequências de recombinação:

Par de genes	Frequência de recombinação (%)
A e *D*	8
B e *D*	13
C e *D*	23

Observe que *C* e *D* exibem a maior porcentagem de recombinação, portanto, *C* e *D* devem estar muito distantes, com os genes *A* e *B* entre eles. Usando as frequências de recombinação e lembrando que 1 m.u. = 1% de recombinação, podemos adicionar *D* a nosso mapa:



Ao fazer vários cruzamentos entre os pares de genes, podemos criar mapas genéticos mostrando a disposição de ligação de vários genes.

Dois pontos devem ser enfatizados sobre construir mapas cromossômicos a partir das frequências de recombinação. Primeiro, lembre-se de que não podemos distinguir entre os genes em diferentes cromossomos e genes localizados distantes no mesmo cromossomo. Se os genes exibem 50% de recombinação, o máximo que podemos dizer sobre eles é que eles pertencem a diferentes grupos de ligação, tanto em diferentes cromossomos, quanto distantes no mesmo cromossomo.

O segundo ponto é que um cruzamento-teste para dois genes que estão distantes no mesmo cromossomo tende a subestimar a distância física real, porque o cruzamento não revela *crossing overs* duplos que podem ocorrer entre os dois genes (**Figura 7.10**). Um *crossing over* duplo surge quando dois eventos de *crossing over* separados ocorrem entre dois *loci*. (No momento, vamos considerar apenas *crossing overs* duplos que ocorrem entre duas das quatro cromátides de um par homólogo – um *crossing over* duplo de duas fitas. Os *crossing overs* duplos que ocorrem entre três de quatro cromátides serão considerados posteriormente, na seção *Efeito dos crossing overs múltiplos*.) Enquanto um *crossing over* único produz combinações de alelos que não estavam presentes nos cromossomos genitores originais, um segundo *crossing over* entre os mesmos dois genes inverte os efeitos do primeiro, restaurando a combinação original dos alelos do genitor (ver Figura 7.10). Portanto, não podemos diferenciar entre os descendentes produzidos pelos *crossing overs* duplos de duas fitas e os descendentes produzidos quando não ocorre nenhum *crossing over*. Como veremos na próxima seção, podemos detectar *crossing overs* duplos se examinarmos um terceiro gene que fica entre os dois *crossing overs*. Como os *crossing overs* duplos entre dois genes não são detectados, as distâncias no mapa serão subestimadas sempre que os *crossing overs* duplos ocorrem. Os *crossing overs* duplos são mais frequentes entre genes que estão distantes, portanto os mapas genéticos que usam curtas distâncias são mais precisos que os que usam distâncias maiores.

Conceitos

Um mapa genético fornece a ordem dos genes em um cromossomo e as distâncias aproximadas de um gene para outro com base nas frequências de recombinação. Nos mapas genéticos, 1% da recombinação é igual a 1 unidade de mapa, ou 1 m.u. Os *crossing overs* duplos entre dois genes passam sem ser detectados, então as distâncias do mapa entre genes distantes tendem a subestimar as distâncias genéticas.

Checagem dos conceitos 3

Qual é a diferença entre um mapa genético e um mapa físico?

Como construir um mapa genético com o uso de cruzamentos-teste de dois pontos

Os mapas genéticos podem ser construídos ao realizar vários cruzamentos-teste. Em cada cruzamento, um dos genitores é

heterozigoto para um par diferente de genes, e as frequências de recombinação são calculadas entre os pares de genes. Um cruzamento-teste entre dois genes é chamado de **cruzamento-teste de dois pontos** ou um cruzamento de dois pontos. Suponha que realizamos vários cruzamentos de dois pontos para quatro genes, *a*, *b*, *c* e *d* e obtemos as seguintes frequências de recombinação:

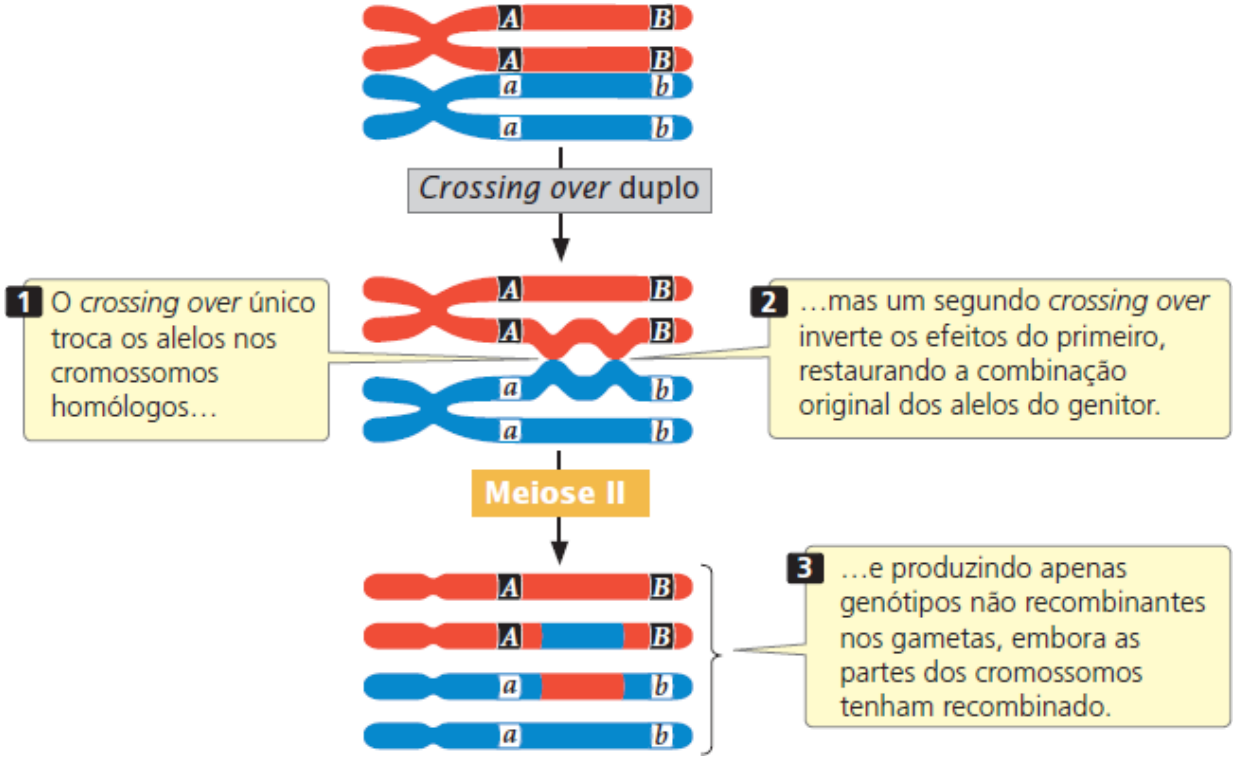
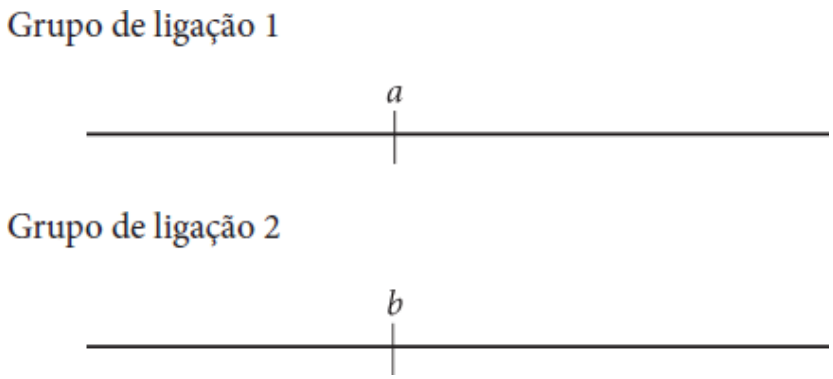


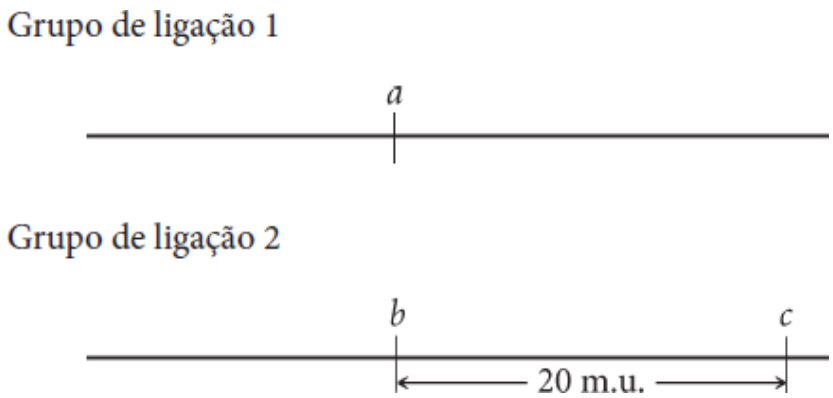
Figura 7.10 Um *crossing over* duplo de duas fitas entre dois genes ligados produz apenas gametas não recombinantes.

***Loci* do gene no cruzamento-teste**	**Frequência de recombinação (%)**
a e *b*	50
a e *c*	50
a e *d*	50
b e *c*	20
b e *d*	10
c e *d*	28

Podemos começar a construir um mapa genético para esses genes ao considerar as frequências de recombinação para cada par. A frequência de recombinação entre *a* e *b* é 50%, que é a frequência de recombinação esperada com a segregação independente. Portanto, os genes *a* e *b* podem estar em cromossomos diferentes ou muito distantes no mesmo cromossomo; os colocaremos em diferentes grupos de ligação com a compreensão de que eles podem estar no mesmo cromossomo ou não:

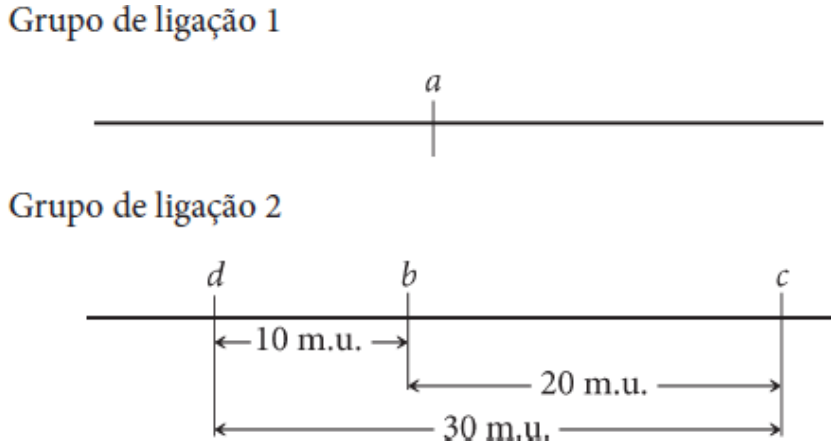


A frequência de recombinação entre *a* e *c* é 50%, indicando que eles também estão em diferentes grupos de ligação. A frequência de recombinação entre *b* e c é 20%, então eles estão ligados e separados por 20 unidades de mapa:



A frequência de recombinação entre *a* e *d* é de 50%, indicando que eles pertencem a diferentes grupos de ligação, enquanto os genes *b* e *d* estão ligados, com uma frequência de recombinação de 10%. Para decidir se o gene *d* está 10 m.u. à esquerda ou à direita do gene *b*, devemos consultar a distância *c*-*d*. Se o gene está 10 m.u. à esquerda do gene *b*, então a distância entre *d* e *c* deve ser aproximadamente a soma da distância entre *b* e *c* e entre *c* e *d*: 20 m.u. + 10 m.u. = 30 m.u. Se, por outro lado, o gene *d* está à direita do gene *b*, então a distância entre os genes *d* e *c* será muito menor, aproximadamente 20 m.u. – 10 m.u. = 10 m.u. As distâncias somadas serão aproximadas porque quaisquer *crossing overs* duplos entre os dois genes estarão ausentes e a distância do mapa será subestimada.

Ao examinar a frequência de recombinação entre *c* e *d*, podemos distinguir entre essas duas possibilidades. A frequência de recombinação entre *c* e *d* é de 28%, então o gene d deve estar à esquerda do gene *b*. Observe que a soma da frequência de recombinação entre *d* e *b* (10%) e entre *b* e *c* (20%) é maior que a frequência de recombinação entre *d* e *c* (28%). Como já foi discutido, essa discrepância surge porque os *crossing overs* duplos entre os dois genes externos passam sem ser detectados, causando uma subestimação da verdadeira distância do mapa. O mapa genético está completo agora:



Resolva o Problema 27

7.3 Um cruzamento-teste de três pontos pode ser usado para mapeamento de três genes ligados

Enquanto os mapas genéticos podem ser construídos a partir de vários cruzamentos-teste para os pares de genes, essa abordagem não é particularmente eficiente porque vários cruzamentos de dois pontos podem ser realizados para estabelecer a ordem dos genes e os *crossing overs* duplos estão ausentes. Uma técnica de mapeamento mais eficiente é um cruzamento-teste para três genes – um **cruzamento-teste de três pontos**, ou cruzamento de três pontos. Com um cruzamento de três pontos, a ordem dos três genes pode ser estabelecida em um único conjunto de descendentes e alguns *crossing overs* duplos podem ser detectados, fornecendo distâncias mais precisas no mapa.

Considere o que acontece quando o *crossing over* ocorre entre três genes hipotéticos ligados. A **Figura 7.11** ilustra um par de cromossomos homólogos de um indivíduo que é heterozigoto em três *loci* (*Aa Bb Cc*). Observe que os genes estão na configuração de acoplamento; todos os alelos dominantes estão em um cromossomo (*A B C*) e todos os alelos recessivos estão em outro (*a b c*). Podem ocorre três tipos de *crossing over* entre esses três genes: dois tipos de *crossing overs* únicos (ver Figura 7.11 A e B) e um *crossing over* duplo (ver Figura 7.11 C). Em cada tipo de *crossing over*, dois dos cromossomos resultantes são recombinantes e dois são não recombinantes.

Observe que nos cromossomos recombinantes resultantes do *crossing over* duplo, os dois alelos externos são os mesmos que nos não recombinantes, mas o alelo do meio é diferente. Esse resultado fornece uma pista importante sobre a ordem dos genes. Nos

descendentes resultantes do *crossing over* duplo, apenas o alelo do meio deve ser diferente dos alelos presentes nos descendentes não recombinantes.

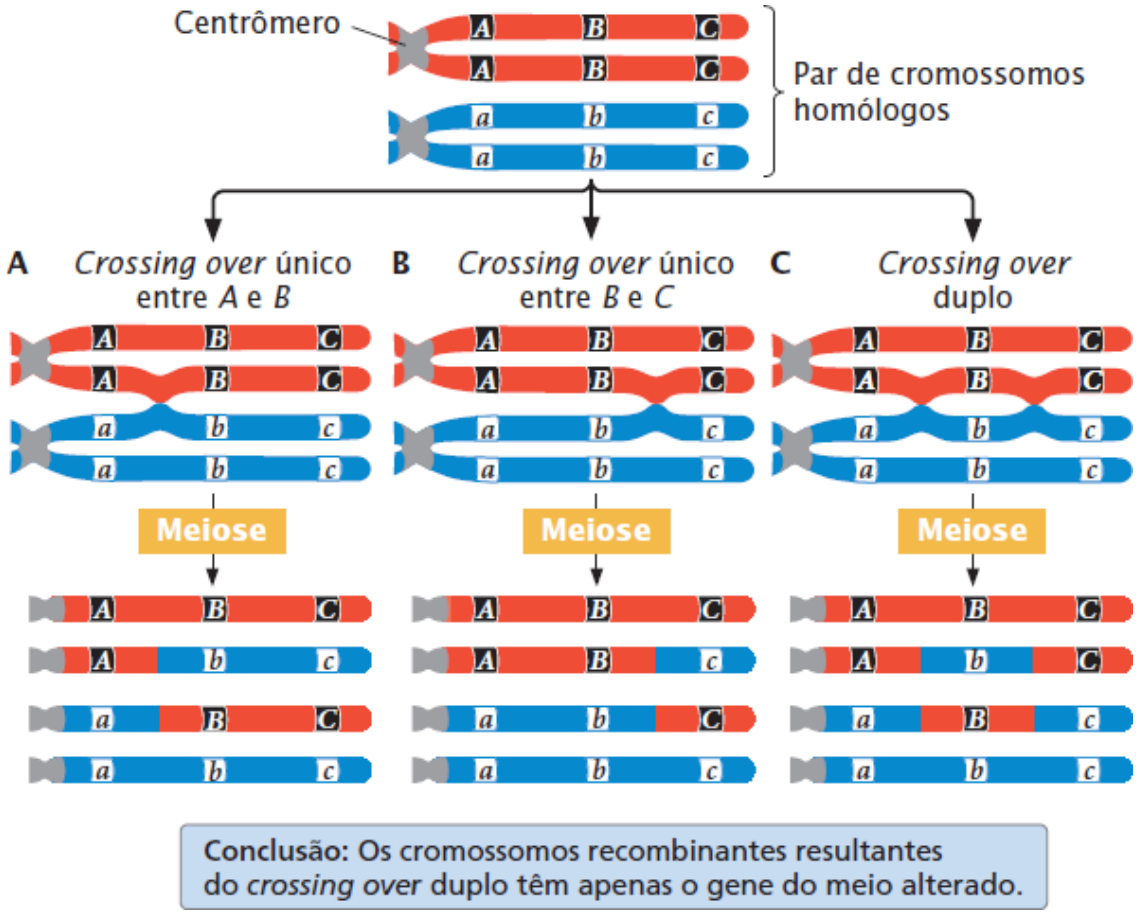


Figura 7.11 Três tipos de *crossing over* que podem ocorrer entre três *loci* ligados.

Como construir um mapa genético com o uso de cruzamento-teste de três pontos

Para examinar o mapeamento genético com cruzamento-teste de três pontos, vamos considerar três mutações recessivas na mosca-da-fruta *Drosophila melanogaster*. Nesta espécie, os olhos escarlate (*st*) são recessivos para os olhos vermelhos tipo selvagem (st^+), a cor ébano do corpo (*e*) é recessiva para a cor cinza tipo selvagem do corpo (e^+), e sem espinha (*ss*) – ou seja, a presença de pequenas cerdas – é recessiva para as cerdas normais tipo selvagem (ss^+). Os *loci* que codificam as três características estão ligados e localizados no cromossomo 3.

Vamos chamar esses três *loci* de *st*, *e* e *ss*, mas tenha em mente que os alelos recessivos (*st*, *e* e *ss*) ou os alelos dominantes (st^+, e^+ e ss^+) podem estar presentes em cada *locus*. Então, quando dizemos que existem 10 m.u. entre *st* e *ss*, queremos dizer que existem 10 m.u. entre os *loci* nos quais as mutações *st* e *ss* ocorrem; também poderíamos dizer que existem 10 m.u. entre st^+ e ss^+.

Para mapear esses genes, precisamos determinar sua ordem no cromossomo e as distâncias genéticas entre eles. Primeiro, vamos montar um cruzamento-teste de três pontos: um cruzamento entre uma mosca heterozigota em todos os três *loci* e uma mosca homozigota para os alelos recessivos em todos os três *loci*. Para produzir moscas heterozigotas para os três *loci*, podemos cruzar várias moscas homozigotas para os alelos do tipo selvagem nos três *loci* com as moscas homozigotas para os alelos recessivos nos três *loci*:

$$\text{P}\quad \frac{st^+\ \ e^+\ \ ss^+}{st^+\ \ e^+\ \ ss^+} \times \frac{st\ \ e\ \ ss}{st\ \ e\ \ ss}$$

$$\downarrow$$

$$\text{F}_1\quad \frac{st^+\ \ e^+\ \ ss^+}{st\ \ e\ \ ss}$$

A ordem dos genes foi arbitrariamente designada porque, nesse momento, não sabemos qual é o gene do meio. Além disso, os alelos nesses heterozigotos estão na configuração de acoplamento (porque todos os alelos dominantes do tipo selvagem foram herdados de um genitor e todas as mutações recessivas de outro genitor), embora o cruzamento-teste também possa ser feito com os alelos em repulsão.

No cruzamento-teste de três pontos, cruzamos os heterozigotos de F_1 com as moscas homozigotas para todas as três mutações recessivas. Em muitos organismos, não faz diferença se o genitor heterozigoto no cruzamento-teste é macho ou fêmea (desde que os genes sejam autossômicos), mas não ocorre *crossing over* nos machos da *Drosophila*. Como o *crossing over* no genitor heterozigoto é essencial para determinar as frequências de recombinação, as moscas heterozigotas no nosso cruzamento-teste devem ser fêmeas. Então, cruzamos as moscas-fêmeas de F_1 que são heterozigotas para todos os três traços com as moscas-machos homozigotas para os traços recessivos:

$$\frac{st^+ \quad e^+ \quad ss^+}{st \quad e \quad ss}\ \text{Fêmea} \times \frac{st \quad e \quad ss}{st \quad e \quad ss}\ \text{Macho}$$

Os descendentes produzidos a partir do cruzamento estão listados na **Figura 7.12**. Para cada *locus*, são produzidas duas classes de descendentes: heterozigotos, apresentando o traço dominante, e homozigotos, apresentando o traço recessivo. Com as duas classes de descendentes possíveis para cada um dos três *loci*, existem $2^3 = 8$ classes de genótipos possíveis nos descendentes. Neste exemplo todas as oito classes fenotípicas estão presentes, mas em alguns dos cruzamentos de três pontos, um ou mais dos fenótipos podem estar ausentes se o número de descendentes for limitado. Contudo, a ausência de uma classe específica pode fornecer informações importantes sobre que combinação de traços é menos frequente e, por fim, sobre a ordem dos genes, como veremos adiante.

Para mapear os genes, precisamos de informações sobre onde e como o *crossing over* ocorreu. No genitor homozigoto recessivo, os dois alelos em cada *locus* são o mesmo, então o *crossing over* não terá efeito nos tipos de gametas produzidos; com ou sem *crossing over*, todos os gametas deste genitor têm um cromossomo com três alelos recessivos (*st* e *ss*). Por outro lado, o genitor heterozigoto tem alelos diferentes em seus dois cromossomos, então o *crossing over* pode ser detectado. As informações que precisamos para o mapeamento, portanto, vêm basicamente dos gametas produzidos pelo genitor heterozigoto. Como os cromossomos contribuídos pelo genitor homozigoto carreiam apenas os alelos recessivos, quaisquer alelos presentes nos cromossomos contribuídos pelo genitor heterozigoto estarão expressos nos descendentes.

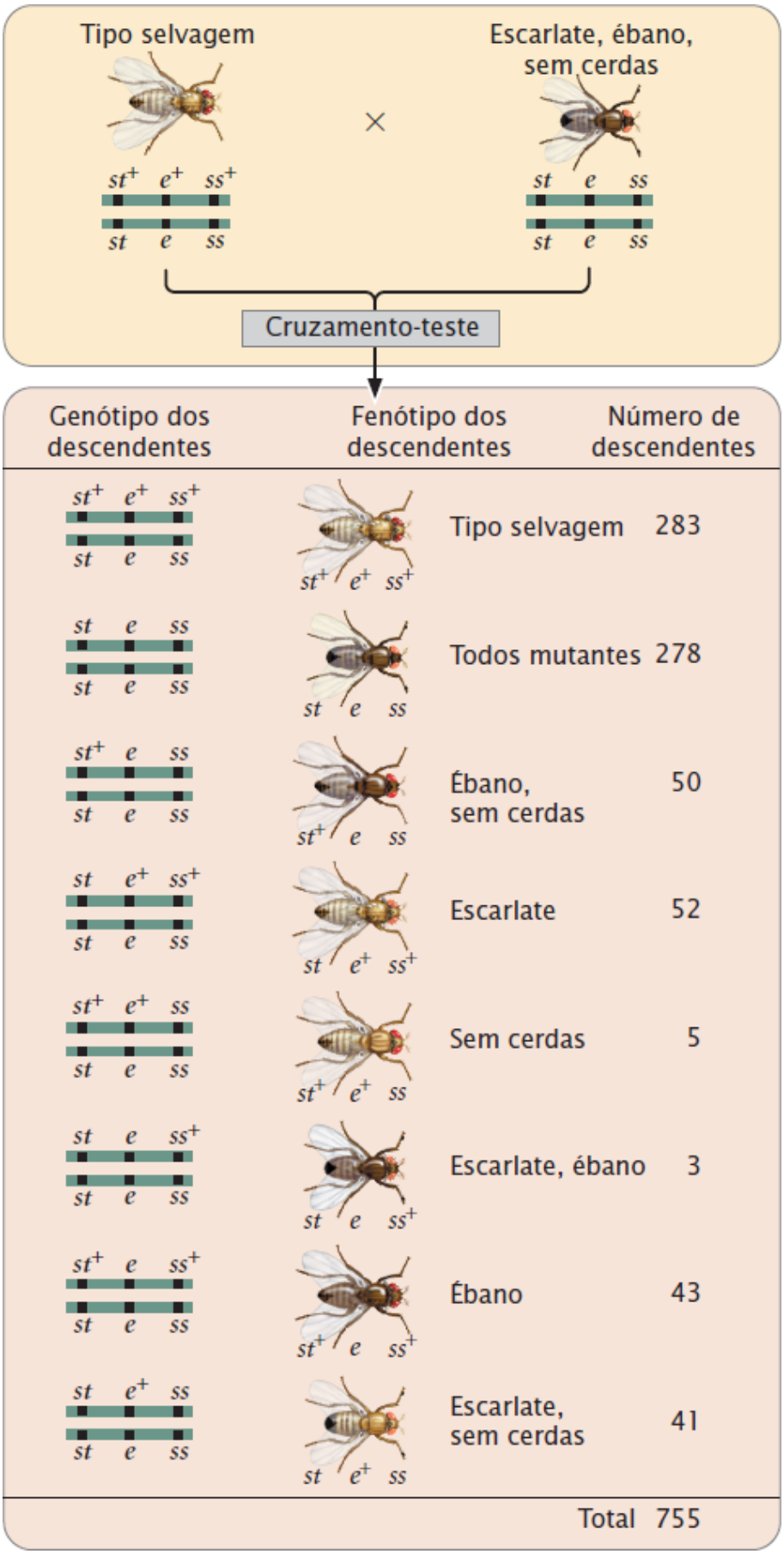


Figura 7.12 Os resultados de um cruzamento-teste de três pontos podem ser usados para mapear os genes ligados. Neste cruzamento-teste de três pontos da *Drosophila melanogaster*, as mutações recessivas de olhos escarlates (*st*), cor de ébano do corpo (*e*) e sem cerdas (*ss*) estão em três *loci* ligados. A ordem dos *loci* foi disposta de forma arbitrária. Cada classe fenotípica inclui as moscas-machos e fêmeas, mas o sexo das moscas descritas é aleatório.

Como um atalho, não escrevemos os genótipos completos dos descendentes do cruzamento-teste, listando apenas os alelos expressos no fenótipo, que são os herdados do genitor heterozigoto. Esta convenção é usada na discussão a seguir.

Conceitos

Para mapear os genes são necessárias as informações sobre localização e número de *crossing overs* nos gametas que produzem os descendentes de um cruzamento. Uma forma eficiente para obter tais informações é usar um cruzamento-teste de três pontos, no qual um heterozigoto individual em três *loci* ligados é cruzado com um homozigoto recessivo nos três *loci*.

Checagem dos conceitos 4

Escreva os genótipos de todos os descendentes recombinantes e não recombinantes esperados a partir do cruzamento de três pontos abaixo:

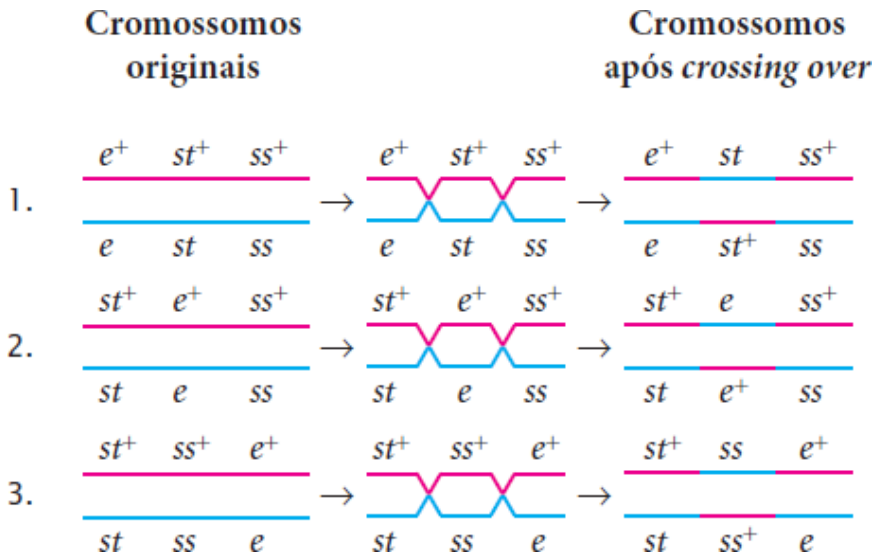
$$\frac{m^+ \quad p^+ \quad s^+}{m \quad p \quad s} \times \frac{m \quad p \quad s}{m \quad p \quad s}$$

Como determinar a ordem dos genes. A primeira tarefa no mapeamento dos genes é determinar sua ordem no cromossomo. Na Figura 7.12, listamos de forma arbitrária os *loci* na ordem *st*, *e*, *ss*, mas não temos como saber qual dos três estava entre os outros dois. Agora, podemos identificar o *locus* do meio ao examinar os descendentes do *crossing over* duplo.

Primeiro, determinamos quais descendentes são não recombinantes; eles serão as duas classes mais numerosas dos descendentes (mesmo se o *crossing over* ocorrer em todas as meioses, os não recombinantes serão pelo menos 50% dos descendentes). Entre os descendentes do cruzamento-teste na Figura 7.12, os mais numerosos são os que têm todos os três traços dominantes (st^+ e^+ ss^+) e aqueles com todos os três traços recessivos (*st e ss*).

A seguir, identificamos os descendentes do *crossing over* duplo. Eles devem ter sempre os dois fenótipos menos numerosos, porque a probabilidade de um *crossing over* duplo é sempre menor que a probabilidade de um *crossing over* único. Os descendentes menos comuns entre os listados na Figura 7.12 são os sem cerdas (st^+ e^+ *ss*) e os com olhos escarlate e corpo cor de ébano (*st e* ss^+), então eles são os descendentes de *crossing over* duplo.

São possíveis três ordens de genes no cromossomo: o *locus* da cor dos olhos poderia estar no meio (*e st ss*), o *locus* da cor do corpo poderia estar no meio (*st e ss*) ou o *locus* da cerda poderia estar no meio (*st ss e*). Para determinar qual gene está no meio, podemos desenhar os cromossomos do genitor heterozigoto com todas as três possíveis ordens de genes e testar se um *crossing over* duplo produz a combinação de genes observada. As três possíveis ordens de genes e os tipos de descendentes produzidos pelos seus *crossing overs* duplos são:



A única ordem de genes que produz cromossomos com o conjunto de alelos observados nos descendentes menos numerosos ou *crossing overs* duplos (st^+ e^+ *ss* e *st e* ss^+ na Figura 7.12) é uma na qual o *locus ss* para as cerdas está no meio (ordem 3 dos genes). Portanto, esta ordem (*st ss e*) deve ser a sequência correta dos genes no cromossomo.

Com um pouco de prática, podemos determinar rapidamente qual *locus* está no meio sem escrever todas as ordens dos genes. Os fenótipos dos descendentes são expressões dos alelos herdados a partir do genitor heterozigoto. Lembre-se de, que quando olhamos os resultados dos *crossing overs* duplos (ver Figura 7.11), apenas os alelos no *locus* do meio são diferentes dos não recombinantes. Se compararmos os descendentes não recombinantes com os do *crossing over* duplo, diferenciamos apenas os alelos no *locus* do meio (**Quadro 7.2**).

Vamos comparar os alelos nos descendentes do *crossing over* duplo st^+ e^+ *ss* com os alelos nos descendentes não recombinantes st^+ e^+ ss^+. Vemos que ambos têm um alelo para olhos vermelhos (st^+) e um para corpo cinza (e^+), mas os não recombinantes têm um alelo para as cerdas normais (ss^+) enquanto os *crossing overs* duplos têm um alelo para cerdas sem espinha (*ss*). Como o *locus* da

cerda é o único diferente, ele tem de estar no meio. Obteremos os mesmos resultados se compararmos outras classes de descendentes de *crossing over* duplo (*st es* s^+) com outra classe de não recombinantes (*st e ss*). Mais uma vez, o único *locus* diferente é o *locus* para cerdas. Não se esqueça de que os não recombinantes e os *crossing overs* duplos têm apenas um *locus* diferente; se eles tiverem dois *loci*, estamos comparando classes erradas de descendentes.

Quadro 7.2 Etapas para determinar a ordem dos genes em um cruzamento de três pontos.

1. Identificar os descendentes não recombinantes (dois genótipos mais numerosos).
2. Identificar os descendentes de *crossing over* duplo (dois genótipos menos numerosos).
3. Comparar o fenótipo dos descendentes do *crossing over* duplo com o fenótipo dos descendentes não recombinantes. Eles devem ser semelhantes em duas características e diferentes em uma.
4. A característica que difere entre os descendentes do *crossing over* duplo e os descendentes não recombinantes é codificada pelo gene do meio.

Conceitos

Para determinar o *locus* médio em um cruzamento de três pontos, compare os descendentes do *crossing over* duplo com os descendentes não recombinantes. Os descendentes do *crossing over* duplo serão as duas classes menos comuns de fenótipos; os não recombinantes serão as duas classes mais comuns de fenótipos. Os descendentes do *crossing over* duplo terão os mesmos alelos que os tipos não recombinantes em dois *loci* e diferentes alelos no *locus* do meio.

Checagem dos conceitos 5

Um cruzamento-teste de três pontos é feito entre três genes ligados. Os descendentes não recombinantes resultantes são s^+ r^+ c^+ e *s r c* e os descendentes do *crossing over* duplo são *s r* c^+ e s^+ r^+ *c*. Qual é o *locus* do meio?

Como determinar as localizações dos crossing overs. Quando sabemos a ordem certa dos *loci* no cromossomo, devemos reescrever os fenótipos dos descendentes do cruzamento-teste na Figura 7.12 com os alelos na ordem correta de modo que possamos determinar onde ocorreram os *crossing overs* (**Figura 7.13**).

Entre as oito classes de descendentes, já identificamos duas classes como não recombinantes (*st*$^+$ *ss*$^+$ *e*$^+$ e *st ss e*) e duas classes de *crossing overs* duplos (*st*$^+$ *ss e*$^+$ e *st ss*$^+$ *e*). As outras quatro classes incluem descendentes que surgiram de um cromossomo que sofreu um *crossing over* único: dois sofreram *crossing overs* únicos entre *st* e *ss* e dois sofreram *crossing overs* únicos entre *ss* e *e*.

Para determinar onde os *crossing overs* ocorreram nesses descendentes, compare os alelos encontrados nos descendentes de *crossing over* único com os encontrados nos não recombinantes, como fizemos para os *crossing overs* duplos. Por exemplo, considere os descendentes com cromossomo *st*$^+$ *ss e*. O primeiro alelo (*st*$^+$) surgiu a partir do cromossomo não recombinante *st*$^+$ *ss*$^+$ *e*$^+$ e os outros dois alelos (*ss* e *e*) devem ter vindo de outro cromossomo não recombinante *st ss e* por meio do *crossing over*:

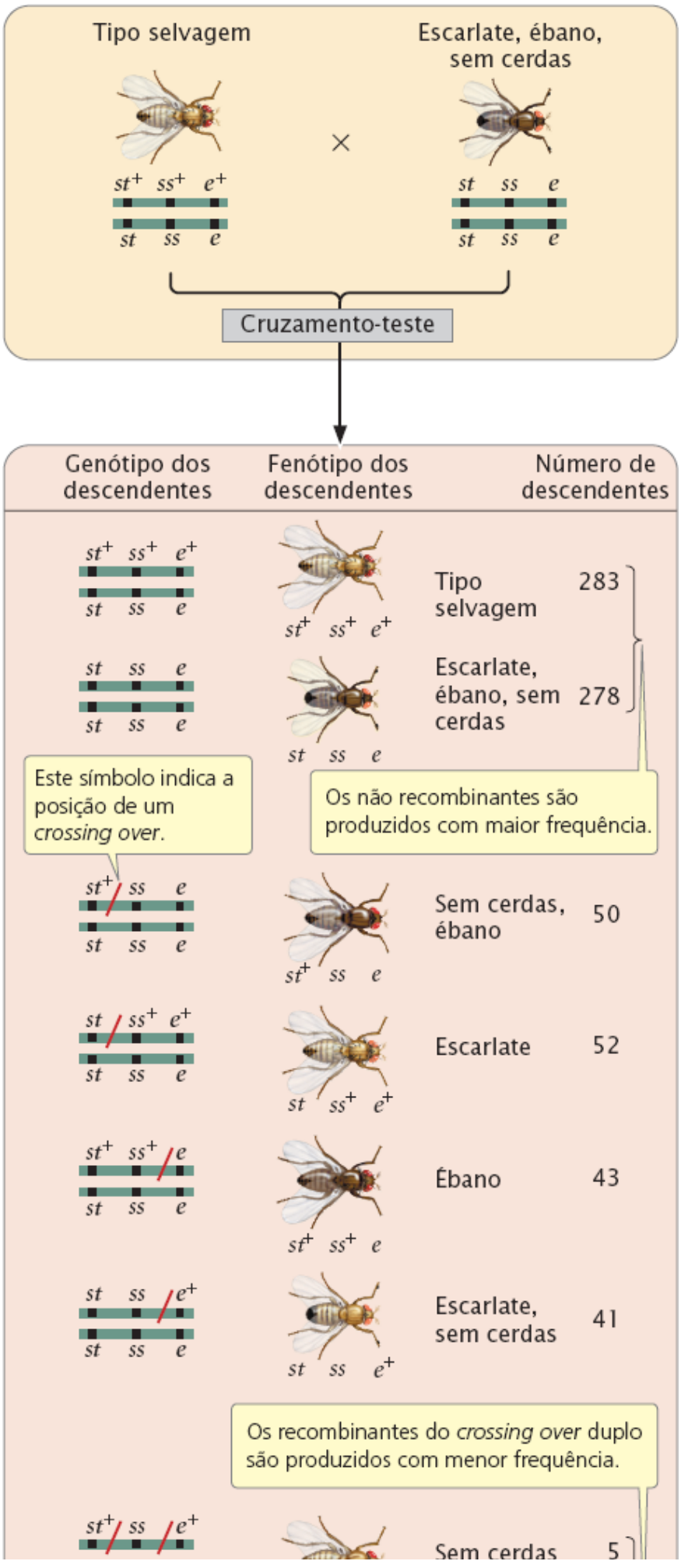
Esse mesmo *crossing over* também produz os descendentes *st ss*$^+$ *e*$^+$.

Esse método também pode ser usado para determinar a localização do *crossing over* nos outros dois tipos de descendentes de *crossing over* único. O *crossing over* entre os cromossomos *ss* e *e* produz *st*$^+$ *ss*$^+$ *e* e *st ss e*$^+$:

Agora sabemos as localizações de todos os *crossing overs*; suas localizações estão marcadas com uma barra na Figura 7.13.

Como calcular as frequências de recombinação. A seguir podemos determinar as distâncias do mapa, que se baseiam nas frequências de recombinação. Calculamos a frequência de recombinação ao somar todos os descendentes recombinantes, dividir este número pelo número total de descendentes do cruzamento e multiplicar o número obtido por 100%. Para determinar as distâncias do

mapa com exatidão, devemos incluir todos os *crossing overs* (únicos e duplos) que ocorreram entre dois genes.



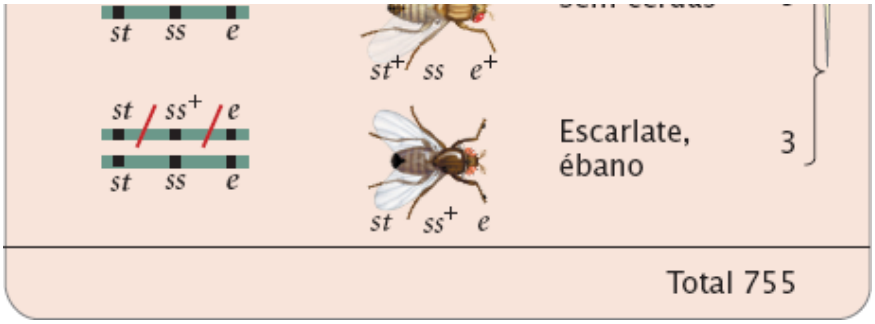


Figura 7.13 Escrever os resultados de um cruzamento-teste de três pontos com os *loci* na ordem correta permite determinar as localizações dos *crossing overs*. Estes resultados são do cruzamento-teste ilustrado na Figura 7.12, com os *loci* apresentados na ordem correta. A localização de um *crossing over* está indicada por uma barra (*/*). Cada classe fenotípica inclui as moscas-machos e fêmeas, mas o sexo das moscas descritas é aleatório.

Os descendentes recombinantes que têm um cromossomo que sofreu *crossing over* entre o *locus* da cor dos olhos (*st*) e o *locus* da cerda (*ss*) incluem os *crossing overs* únicos (*st*$^+$ / *ss e* e *st* / *ss*$^+$ *e*$^+$) e os *crossing overs* duplos (*st*$^+$ / *ss* / *e*); ver Figura 7.13. Existe um total de 755 descendentes, então a frequência de recombinação entre *ss* e *st* é:

$$\text{Frequência de recombinação de } st\text{–}ss = \frac{50 + 52 + 5 + 3}{775} \times 100\% = 14{,}6\%$$

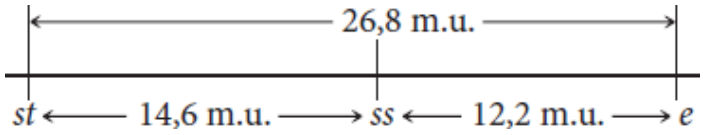
A distância entre os *loci st* e *ss* pode ser expressa como 14,6 m.u.

A distância de mapa entre o *locus* de cerda (*ss*) e o *locus* do corpo (*e*) é determinada da mesma forma. Os descendentes recombinantes com um *crossing over* entre *ss* e *e* são os *crossing overs* únicos *st*$^+$ *ss*$^+$ */e* e *st ss*/*e*$^\pm$ e os *crossing overs* duplos *st*$^+$ / *ss*/ *e*$^+$ e *st* / *ss*$^+$ / *e*. A frequência de recombinação é:

$$\text{Frequência de recombinação de } ss\text{–}e = \frac{43 + 41 + 5 + 3}{755} \times 100\% = 12{,}2\%$$

Assim, a distância de mapa entre *ss* e *e* é 12,2 m.u.

Finalmente, calculamos a distância de mapa entre os outros dois *loci st* e *e*. Esta distância de mapa pode ser obtida ao somar as distâncias entre *st* e *ss* e entre *ss* e *e* (14,6 m.u. + 12,2 m.u. = 26,8 m.u.). Agora podemos usar as distâncias de mapa para desenhar um mapa dos três genes no cromossomo:



Um mapa genético da *D. melanogaster* está ilustrado na **Figura 7.14**.

Interferência e coeficiente de coincidência. Além de dar informações sobre as distâncias que separam os genes, as distâncias de mapa também indicam as proporções de gametas recombinantes e não recombinantes que serão produzidos em um cruzamento. Por exemplo, saber que os genes *st* e *ss* no terceiro cromossomo de *D. melanogaster* são separados por uma distância de 14,6 m.u. indica que 14,6% dos gametas produzidos por uma mosca heterozigota nestes dois *loci* serão recombinantes. Da mesma forma, 12,2% dos gametas de uma mosca heterozigota para *ss* e *e* serão recombinantes.

Teoricamente, devemos ser capazes de calcular a proporção de gametas duplo-recombinantes ao usar a regra de multiplicação de probabilidade (ver *Capítulo 3*). Ao aplicá-la, descobriremos que a proporção (probabilidade) dos gametas com *crossing overs* duplos entre *st* e *e* é igual à probabilidade de recombinação entre *st* e *ss* multiplicada pela probabilidade de recombinação entre *ss* e *e*, ou 0,146 × 0,122 = 0,0178. Multiplicar esta probabilidade pelo número total de descendentes nos indica o número esperado de descendentes de *crossing over* duplo a partir do cruzamento: 0,0178 × 755 = 13,4. Apenas 8 *crossing overs* duplos – consideravelmente menos que os 13 esperados – foram observados nos descendentes do cruzamento (ver Figura 7.13).

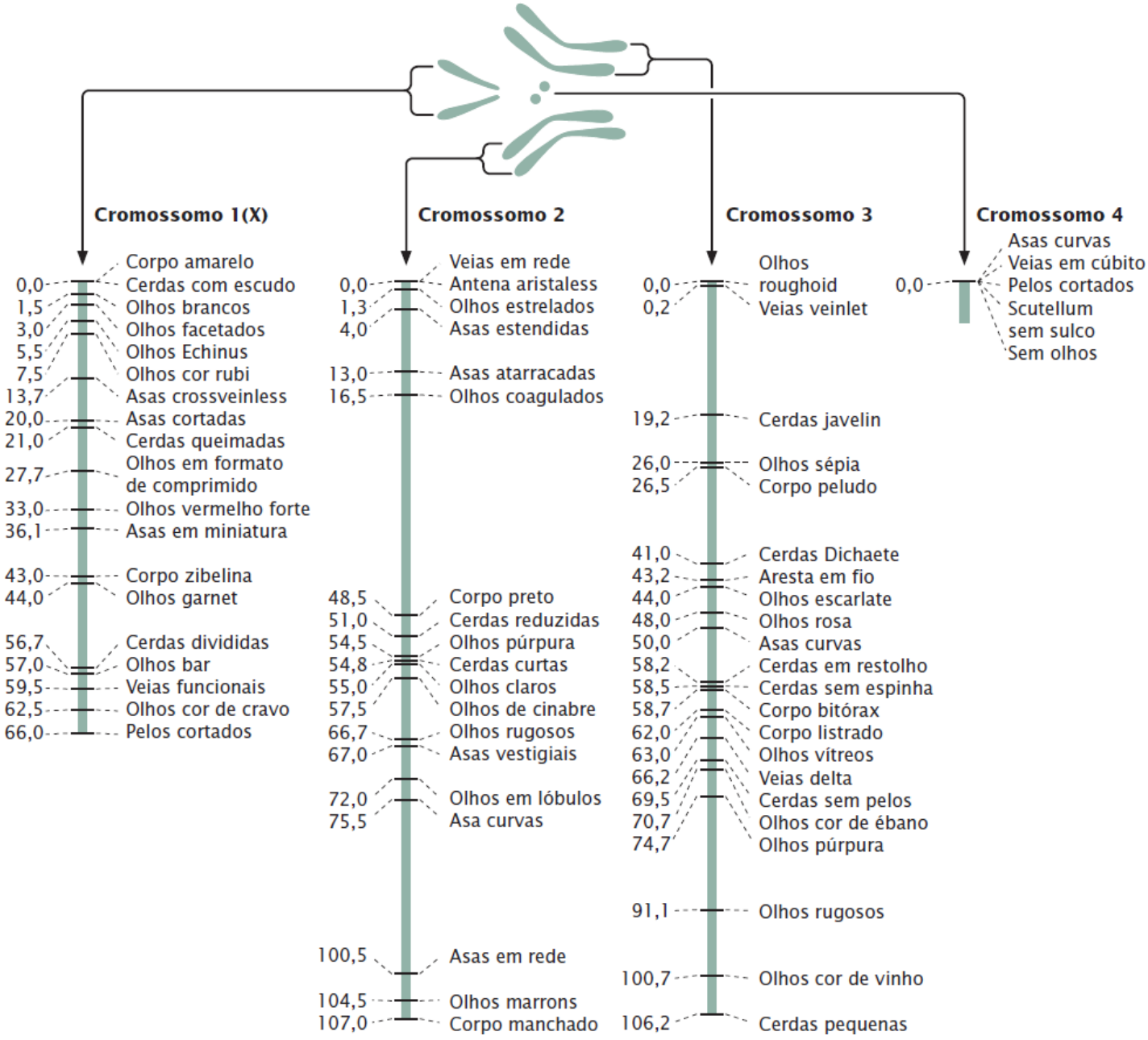


Figura 7.14 A *Drosophila melanogaster* tem quatro grupos de ligação correspondentes aos seus quatro pares de cromossomos. Esses genes foram mapeados usando as frequências de recombinação. As distâncias entre os genes de um grupo de ligação estão em unidades de mapa.

Esse fenômeno é comum nos organismos eucarióticos. O cálculo supõe que cada evento de *crossing over* é independente e que a ocorrência de um *crossing over* não influencia a ocorrência de outro. Mas os *crossing overs* frequentemente *não* são eventos independentes: a ocorrência de um *crossing over* tende a inibir *crossing overs* adicionais na mesma região do cromossomo, de modo que os *crossing overs* duplos são menos frequentes que o esperado.

O grau de influência de um *crossing over* em outros na mesma região é chamado de **interferência**. Para calcular a interferência, primeiro determinamos o **coeficiente de coincidência**, que é a razão dos *crossing overs* duplos observados para os *crossing overs* duplos esperados:

$$\text{Coeficiente de coincidência} = \frac{\text{número de } \textit{crossing overs} \text{ duplos observados}}{\text{número de } \textit{crossing overs} \text{ duplos esperados}}$$

Para os *loci* mapeados no terceiro cromossomo da *D. melanogaster* (ver Figura 7.13), descobrimos que

$$\text{Coeficiente de coincidência} = \frac{5 + 3}{0{,}146 \times 0{,}122 \times 755} = \frac{8}{13{,}4} = 0{,}6$$

que indica que estamos de fato observando apenas 60% de *crossing overs* duplos que esperamos com base nas frequências de *crossing over* único. A interferência é calculada como:

$$\text{interferência} = 1 - \text{coeficiente de coincidência}$$

Então, a interferência para nosso cruzamento de três pontos é:

$$\text{interferência} = 1 - 0{,}6 = 0{,}4$$

Este valor de interferência significa que 40% dos descendentes do *crossing over* duplo esperado não serão observados, por causa da interferência. Quando a interferência é completa e não são observados descendentes do *crossing over* duplo, o coeficiente de coincidência é 0 e a interferência é 1.

Em alguns casos, um *crossing over aumenta* a probabilidade de que outro *crossing over* ocorra próximo e então vemos *mais* descendentes de *crossing over* duplo que esperávamos. Neste caso, o coeficiente de coincidência é maior que 1 e a interferência é negativa.

A maioria dos organismos eucarióticos exibe interferência, que faz com que os *crossing overs* sejam mais espaçados que o esperado de forma aleatória. A interferência foi observada pela primeira vez nos cruzamentos de *Drosophila* no início de 1900 e, até o momento, apesar de anos de estudo, o mecanismo pelo qual a interferência ocorre ainda não está bem compreendido. Um modelo proposto de interferência sugere que os *crossing overs* ocorrem quando se acumula tensão ao longo do cromossomo. De acordo com este modelo, um *crossing over* libera tensão por alguma distância ao longo do cromossomo. Como o *crossing over* libera a tensão que provoca *crossing overs*, é menos provável que ocorram *crossing overs* adicionais na mesma área. **Resolva o Problema 29**

Conceitos

O coeficiente de coincidência é igual ao número de *crossing overs* duplos observados divididos pelo número de *crossing overs* duplos esperados com base nas frequências de *crossing overs* únicos. A interferência é igual a 1 – coeficiente de coincidência; ela indica o grau no qual um *crossing over* interfere com *crossing overs* adicionais.

Checagem dos conceitos 6

Ao analisar os resultados de um cruzamento-teste de três pontos, um estudante determina que a interferência é –0,23. O que este valor negativo de interferência indica?

a. Ocorrem menos *crossing overs* duplos que o esperado com base nas frequências de *crossing overs* únicos.
b. Ocorrem mais *crossing overs* duplos que o esperado com base nas frequências de *crossing overs* únicos.
c. Ocorre menos *crossing overs* únicos que o esperado.
d. Um *crossing over* em uma região interfere com *crossing overs* adicionais na mesma região.

Conceitos conectantes

Como percorrer o cruzamento de três pontos

Agora examinamos o cruzamento de três pontos em detalhes e vimos como as informações derivadas do cruzamento podem ser usadas para mapear uma série de três genes ligados. Vamos revisar rapidamente as etapas necessárias para mapear genes a partir de um cruzamento de três pontos.

1. **Escreva os fenótipos e números de descendentes produzidos no cruzamento de três pontos.** É mais fácil interpretar os fenótipos dos descendentes se você usar símbolos alélicos para os traços (como st^+ e^+ *ss*).
2. **Escreva os genótipos dos genitores originais usados para produzir o indivíduo heterozigoto triplo** em um cruzamento-teste e, se conhecida, a disposição (acoplamento ou repulsão) dos alelos nos seus cromossomos.
3. **Determine quais classes fenotípicas entre os descendentes são as não recombinantes e quais são as de *crossing overs* duplos.** Os não recombinantes serão os dois fenótipos mais comuns, os *crossing overs* duplos serão os dois fenótipos menos comuns.
4. **Determine qual *locus* está no meio.** Compare os alelos presentes nos *crossing overs* duplos com os presentes nos não recombinantes; cada classe de *crossing overs* duplos deve ser como um dos não recombinantes para dois *loci* e devem ser diferentes para um *locus*. O *locus* diferente é o do meio.
5. **Reescreva os fenótipos com os genes na ordem correta.**
6. **Determine onde os *crossing overs* ocorreram para dar origem aos fenótipos dos descendentes.** Para isso, compare cada fenótipo com o fenótipo dos descendentes não recombinantes.
7. **Determine as frequências de recombinação.** Adicione os números de descendentes que têm um cromossomo com um *crossing over* entre um par de *locus*. Adicione os *crossing overs* duplos a esse número.

Divida a soma pelo número total de descendentes do cruzamento e multiplique por 100%; o resultado é a frequência de recombinação entre os *loci*, que é a mesma que a distância do mapa.

8. **Desenhe um mapa dos três *loci*.** Indique qual *locus* está no meio e indique as distâncias entre eles.
9. **Determine o coeficiente de coincidência e a interferência.** O coeficiente de coincidência é o número de descendentes do *crossing over* duplo observado dividido pelo número de descendentes do *crossing over* duplo esperado. O número esperado pode ser obtido ao multiplicar o produto pelas duas probabilidades de recombinação única pelo número total de descendentes no cruzamento.

Problema desenvolvido

Na *D. melanogaster*, as asas de querubim (*ch*), corpo preto (*b*) e olhos cor de cinabre (*cn*) surgem a partir de alelos recessivos que estão localizados no cromossomo 2. Uma mosca homozigota do tipo selvagem foi cruzada com uma mosca com asas de querubim, corpo preto e olhos cor de cinabre, e foi feito o cruzamento-teste das fêmeas F_1 resultantes com machos com asas de querubim, corpo preto e olhos cor de cinabre. Foram produzidos os seguintes descendentes a partir do cruzamento-teste:

ch	b^+	*cn*	105
ch^+	b^+	cn^+	750
ch^+	*b*	*cn*	40
ch^+	b^+	*cn*	4
ch	*b*	*cn*	753
ch	b^+	cn^+	41
ch^+	*b*	cn^+	102
ch	*b*	cn^+	5
Total			1.800

a. Determine a ordem linear dos genes no cromossomo. Qual gene está no meio?)
b. Calcule as distâncias recombinantes entre os três *loci*.
c. Determine o coeficiente de coincidência e a interferência para os três *loci*.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

A ordem dos genes no cromossomo, as distâncias recombinantes entre os genes, o coeficiente de coincidência e a interferência.

Quais informações são fornecidas para solucionar o problema?

- Uma mosca homozigota do tipo selvagem foi cruzada com uma mosca com asas de querubim, corpo preto e olhos cor de cinabre e foi feito o cruzamento-teste das fêmeas F_1 resultantes com machos com asas de querubim, corpo preto e olhos cor de cinabre
- Os números dos diferentes tipos de moscas que surgem entre os descendentes do cruzamento-teste.

Etapas para a solução

a. Podemos representar os cruzamentos neste problema como o esquema a seguir:

$$\text{P} \quad \frac{ch^+ \quad b^+ \quad cn^+}{ch^+ \quad b^+ \quad cn^+} \times \frac{ch \quad b \quad cn}{ch \quad b \quad cn}$$

$$\downarrow$$

$$\text{F}_1 \quad \frac{ch^+ \quad b^+ \quad cn^+}{ch \quad b \quad cn}$$

$$\text{Cruzamento-teste} \quad \frac{ch^+ \quad b^+ \quad cn^+}{ch \quad b \quad cn} \times \frac{ch \quad b \quad cn}{ch \quad b \quad cn}$$

Observe que neste ponto não sabemos a ordem dos genes e colocamos *b* arbitrariamente no meio.

A próxima etapa é determinar quais dos descendentes do cruzamento-teste são não recombinantes e quais são *crossing overs* duplos. Os não recombinantes devem ser o fenótipo mais frequente, então eles têm de ser os descendentes com fenótipos codificados por $\underline{ch^+ \; b^+ \; cn^+}$ e $\underline{ch \; b \; cn}$. Esses genótipos estão em acordo com os dos genitores, indicados anteriormente. Os *crossing*

overs duplos são os fenótipos menos frequentes e estão codificados por *ch b cn* e *ch b cn* .

Podemos determinar a ordem dos genes ao comparar os alelos presentes nos *crossing overs* duplos com os presentes nos não recombinantes. Os descendentes do *crossing over* duplo devem ser semelhantes a um dos não recombinantes em dois *loci* e diferentes em um *locus*; o alelo diferente deve estar no meio. Compare os descendentes *ch b* cn^+ do *crossing over* duplo com *c b cn* não recombinante. Ambos têm asas de querubim (*ch*) e corpo preto (*b*), mas os descendentes do *crossing over* duplo têm olhos do tipo selvagem (cn^+), enquanto os não recombinantes têm olhos cor de cinabre (*cn*). O *locus* que determina os olhos cor de cinabre deve estar no meio.

b. Para calcular as frequências de recombinações entre os genes, primeiro escrevemos os fenótipos dos descendentes com os genes que os codificam na ordem certa. Já identificamos os descendentes não recombinantes e do *crossing over* duplo; então, os outros quatro tipos de descendentes devem ser resultado de *crossing overs* únicos. Para determinar onde os *crossing overs* únicos ocorreram, comparamos os alelos encontrados nos descendentes do *crossing over* único com os alelos dos não recombinantes. O *crossing over* tem de ocorrer onde os alelos são trocados daqueles encontrados em um não recombinante para os encontrados em outro não recombinante. As localizações dos *crossing overs* estão indicadas por uma barra:

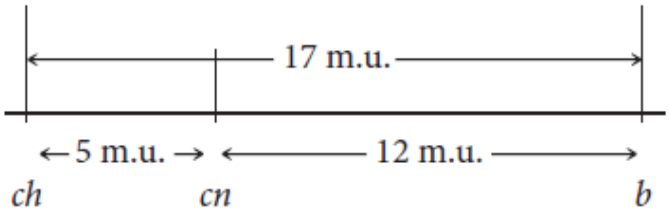
ch		*cn*	/	b^+	105	*crossing over* único
ch^+		cn^+		b^+	750	não recombinante
ch^+	/	*cn*		*b*	40	*crossing over* único
ch^+	/	*cn*	/	b^+	4	*crossing over* duplo
ch		*cn*		*b*	753	não recombinante
ch	/	cn^+		b^+	41	*crossing over* único
ch^+		cn^+	/	*b*	102	*crossing over* único
ch	/	cn^+	/	*b*	5	*crossing over* duplo
Total					1.800	

A seguir, determinamos as frequências de recombinação e desenhamos um mapa genético:

$$\text{frequência de recombinação de } ch\text{-}cn = \frac{40 + 4 + 41 + 5}{1.800} \times 100\% = 5\%$$

$$\text{frequência de recombinação de } cn\text{-}b = \frac{105 + 4 + 102 + 5}{1.800} \times 100\% = 12\%$$

$$\text{distância de mapa } ch\text{-}b = 5\% = 12\% = 17\%$$



c. O coeficiente de coincidência é o número de *crossing overs* duplos observados dividido pelo número de *crossing overs* duplos esperados. O número de *crossing overs* duplos esperados é obtido ao multiplicar a probabilidade de um cruzamento entre *ch* e *cn* (0,05) × a probabilidade de um *crossing over* entre *cn* e *b* (0,12) × o número total de descendentes no cruzamento (1.800):

$$\text{Coeficiente de coincidência} = \frac{4 + 5}{0{,}05 \times 0{,}12 \times 1.800} = 0{,}83$$

Finalmente, a interferência é igual a 1 – o coeficiente de coincidência:

$$\text{interferência} = 1 - 0{,}83 = 0{,}17$$

Para melhorar sua habilidade com os cruzamentos de três pontos, resolva o **Problema 30** no final deste capítulo.

Efeito dos *crossing overs* múltiplos

Até o momento, examinamos os efeitos dos *crossing overs* duplos que ocorrem entre apenas duas das quatro cromátides (fitas) de um par homólogo. Eles são chamados de *crossing overs* de duas fitas. Os *crossing overs* duplos incluindo três e até quatro das cromátides de um par homólogo também podem ocorrer (**Figura 7.15**). Se examinarmos apenas os alelos nos *loci* em um lado de ambos os eventos de *crossing over*, os *crossing overs* duplos de duas fitas não resultam em novas recombinações de alelos e não são produzidos gametas recombinantes (ver Figura 7.15). Os *crossing overs* duplos de três fitas resultam em dois dos quatros gametas sendo recombinantes e os de quatro fitas geram todos os quatro gametas sendo recombinantes. Assim, os *crossing overs* duplos de duas fitas produzem 0% de recombinação, os *crossing overs* duplos de três fitas produzem 50% de recombinação e os *crossing overs* duplos de quatro fitas produzem 100% de recombinação. O resultado geral é que todos os tipos de *crossing overs* duplos, em conjunto, produzem uma média de 50% de descendentes recombinantes.

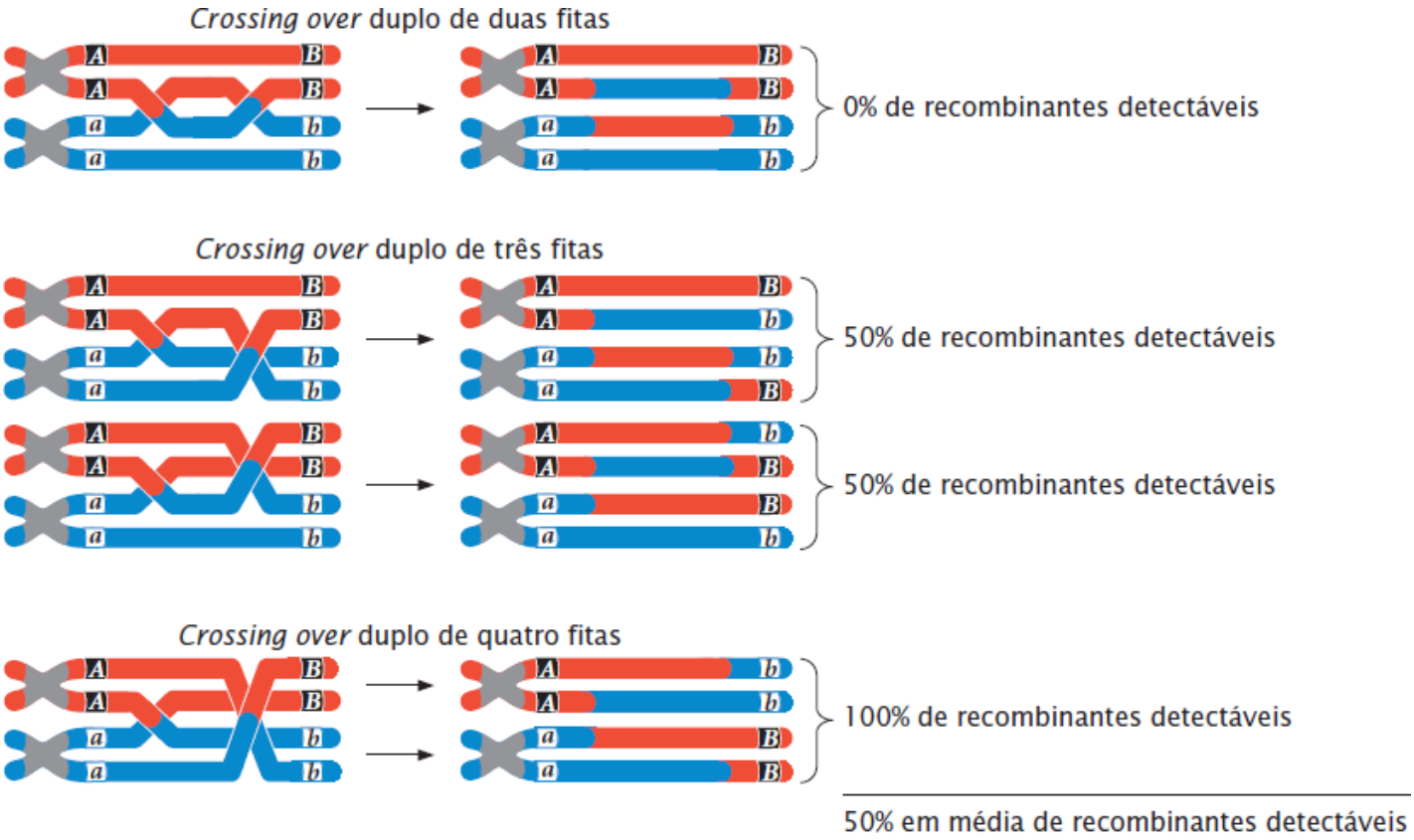


Figura 7.15 Resultados de *crossing overs* duplos de duas, três e quatro fitas na recombinação entre dois genes.

Como observamos, os *crossing overs* duplos de duas fitas fazem com que os alelos de ambos os lados dos *crossing overs* permaneçam o mesmo e não produzam descendentes recombinantes. Os *crossing overs* de três e quatro fitas produzem descendentes recombinantes, mas são os mesmos tipos que os produzidos por *crossing overs* únicos. Consequentemente, alguns *crossing overs* múltiplos não são detectados quando são observados os descendentes de um cruzamento genético. Portanto, as distâncias do mapa com base nas taxas de recombinações subestimam as reais distâncias físicas entre os genes porque alguns *crossing overs* múltiplos não são detectados entre os descendentes de um cruzamento. Quando os genes estão muito próximos, *crossing overs* múltiplos são improváveis e as distâncias com base em taxas de recombinação correspondem de forma exata às distâncias físicas no cromossomo. Mas, à medida que a distância entre os genes aumenta, mais *crossing overs* múltiplos são prováveis e a discrepância entre as distâncias genéticas (com base nas taxas de recombinação) e distâncias físicas aumenta. Para corrigir essa discrepância, os geneticistas desenvolveram **funções de mapeamento** matemáticas, que relacionam as frequências de recombinação às distâncias físicas reais entre os genes (**Figura 7.16**). A maioria dessas funções se baseia na distribuição de Poisson, que prevê a probabilidade de eventos raros múltiplos. Com o uso de tais funções de mapeamento, as distâncias do mapa com base nas taxas de recombinação podem ser estimadas de forma mais precisa.

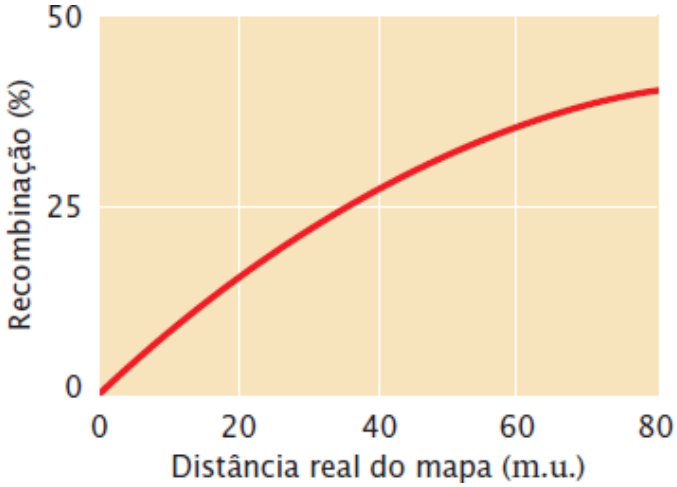


Figura 7.16 **A recombinação percentual subestima a distância física real entre os genes em distâncias maiores do mapa.**

Como mapear os genes humanos

Os esforços no mapeamento dos genes humanos são prejudicados pela incapacidade de realizar os cruzamentos desejados e o número reduzido de descendentes na maioria das famílias. Os geneticistas ficam restritos a análises de heredogramas, que na maioria das vezes são incompletos e fornecem informações limitadas. Contudo, um grande número de traços humanos foi mapeado com sucesso com o uso de dados de heredograma para analisar as ligações. Como o número de descendentes de qualquer acasalamento em geral é pequeno, os dados de várias famílias e heredogramas são combinados para testar a segregação independente. Os métodos usados nesses tipos de análise são complexos, mas um exemplo ilustrará como a ligação pode ser detectada a partir de dados de heredograma.

Uma das primeiras demonstrações documentadas de ligação nos seres humanos foi entre o *locus* da síndrome da unha-patela e o *locus* que determina os tipos sanguíneos do sistema AB0. A síndrome da unha-patela é um distúrbio autossômico dominante caracterizado por unhas anormais e ausência ou atrofia das patelas. Os tipos sanguíneos do sistema AB0 são determinados por um *locus* autossômico com alelos múltiplos (ver *Capítulo 5*). A ligação entre os genes que codificam esses traços foi estabelecida em famílias nas quais ambos os traços se separam. Parte de uma família está ilustrada na **Figura 7.17**.

A síndrome da unha-patela é rara; então podemos assumir que as pessoas que têm este traço são heterozigotas (*Nn*) e as não afetadas são homozigotas (*nn*). Os genótipos AB0 podem ser deduzidos a partir dos fenótipos e dos tipos de descendentes produzidos. A pessoa I-2 na Figura 7.17, por exemplo, tem tipo sanguíneo B, com dois possíveis genótipos: I^BI^B ou I^Bi (ver Figura 5.6). Como alguns de seus descendentes têm tipo sanguíneo 0 (genótipo *ii*) e, portanto, herdaram obrigatoriamente um alelo *i* de cada genitor, a mulher I-2 tem de apresentar o genótipo I^Bi. Da mesma maneira, a ocorrência de descendentes com tipo sanguíneo 0 na geração II indica que o homem I-1, com tipo sanguíneo A, também tinha de ser portador de um alelo *i* e, portanto, tem genótipo I^Ai. Os pais desta família são:

$$I^Ai\ Nn \times I^Bi\ nn$$

A partir da geração II, podemos ver que os genes para a síndrome da unha-patela e os tipos sanguíneos não se separam de modo independente. Todas as crianças na geração II com síndrome da unha-patela tinham tipo sanguíneo B ou 0; todas as crianças com tipo sanguíneo A tinham unhas e patelas normais. Se os genes que a síndrome da unha-patela e os tipos sanguíneos AB0 se separassem de forma independente, seria esperado que algumas crianças na geração II tivessem tipo sanguíneo A e síndrome da unha-patela, herdando ambos os alelos I^A e *N* de seu pai. Este desfecho indica que as disposições dos alelos nos cromossomos nos pais são:

$$\frac{I^A \quad n}{i \quad N} \times \frac{I^B \quad n}{i \quad n}$$

O heredograma indica que não há recombinação entre os descendentes (geração II) desses pais, mas existem dois casos de recombinação entre as pessoas na geração III. As pessoas II-1 e II-2 têm os seguintes genótipos:

$$\frac{i^B \quad n}{i \quad N} \times \frac{I^A \quad n}{i \quad n}$$

A criança III-2 tem o tipo sanguíneo A e não tem a síndrome de unha-patela; então ela necessariamente tem o genótipo

$$\frac{I^A \quad n}{i \quad n}$$

e deve ter herdado ambos os alelos *i* e *n* de seu pai. Estes alelos estão em diferentes cromossomos no pai, então ocorre *crossing over*. O *crossing over* também ocorreu para gerar a criança III-3.

No heredograma da Figura 7.17, treze crianças de acasalamentos nos quais os genes que codificam a síndrome da unha-patela e os tipos sanguíneos AB0 se separam, dois deles são recombinantes. Assim, assumimos que os *loci* para esta síndrome e os tipos sanguíneos AB0 estão ligados, com uma frequência de recombinação de 2/13 = 0,154. Entretanto, é possível que os genes estejam se separando de forma independente e que o pequeno número de crianças dê a impressão de que os genes estão ligados. Para determinar a probabilidade de que os genes estejam de fato ligados, os geneticistas calculam os ***lod scores*** (logaritmo de probabilidade).

Para obter um *lod score*, calculamos a probabilidade de se obterem os resultados observados com a suposição de que os genes estão ligados em um grau específico de recombinação, e com a suposição de segregação independente. Então, determinamos a razão dessas duas probabilidades e o logaritmo dessa razão é o *lod score*. Suponha que a probabilidade de se obter um conjunto específico de observações com a hipótese de ligação e uma frequência de recombinação é de 0,1 e que a probabilidade de obter as mesmas observações com a hipótese de segregação independente é de 0,0001. A razão dessas duas probabilidades é 0,1/0,0001 = 1.000, cujo logaritmo (*lod score*) é 3. Assim, a ligação com uma recombinação específica é 1.000 vezes tão possível quanto uma segregação independente para produzir o que foi observado. Um *lod score* de 3 ou mais, em geral, é considerado um indício convincente para ligação. **Resolva o Problema 36**

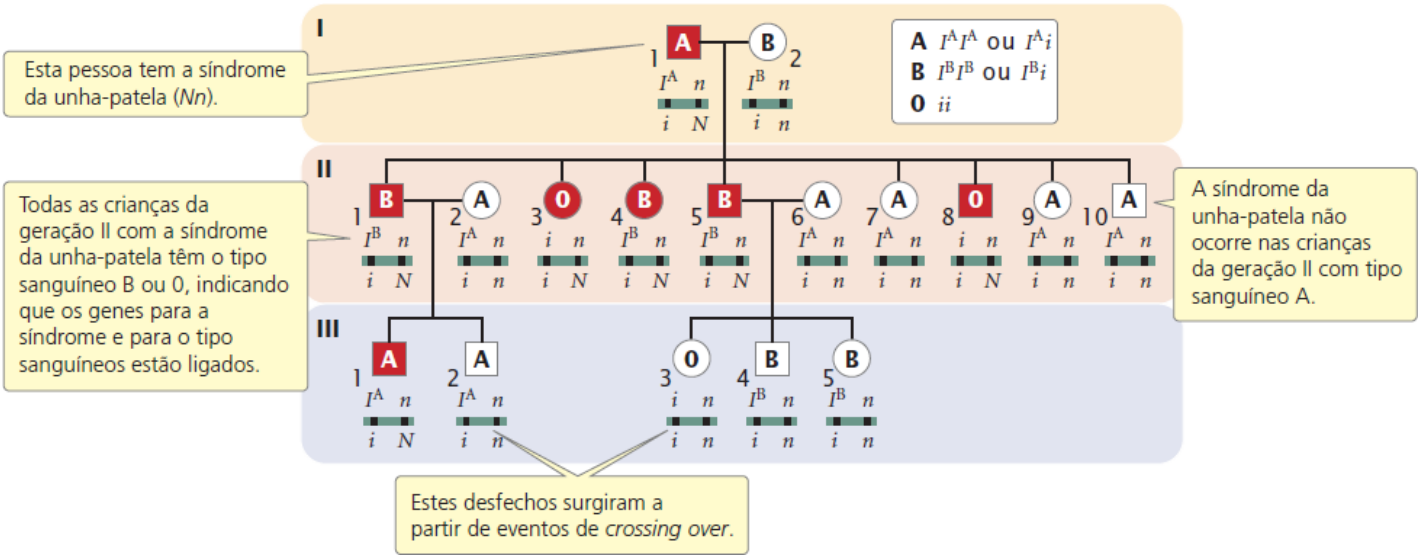


Figura 7.17 A ligação entre os tipos sanguíneos do sistema AB0 e a síndrome da unha-patela foi estabelecida ao examinar famílias nas quais ambos os traços se separam. O heredograma apresentado aqui é para uma dessas famílias. O tipo sanguíneo AB0 está indicado em cada círculo ou quadrado. O genótipo, deduzido a partir do fenótipo, é fornecido abaixo de cada círculo ou quadrado.

Mapeamento com marcadores moleculares

Por muitos anos, o mapeamento dos genes estava limitado na maioria dos organismos por causa da disponibilidade dos **marcadores genéticos** – genes variáveis com fenótipos de fácil visualização para os quais a herança poderia ser estudada. Os marcadores genéticos tradicionais incluem genes que codificam características de fácil observação, como cor da flor, formato da semente, tipos sanguíneos ou diferenças bioquímicas. A escassez desses tipos de características em muitos organismos limitou as tentativas de mapeamento.

Nos anos de 1980, novas técnicas moleculares possibilitaram examinar variações no próprio DNA, fornecendo um número quase ilimitado de marcadores genéticos que podem ser usados para criar mapas genéticos e estudar as relações de ligação. Os primeiros desses marcadores eram polimorfismos de comprimento de fragmentos de restrição (RFLPs), que são variações na sequência de DNA detectadas ao cortar o DNA com enzimas de restrição (ver *Capítulo 19*). Posteriormente, foram desenvolvidos métodos para detectar muitas variáveis de sequências pequenas de DNA repetidas uma após a outra, chamadas de microssatélites. Agora, o sequenciamento de DNA permite a detecção direta de variações individuais nos nucleotídios do DNA. Todos esses métodos expandiram a disponibilidade de marcadores genéticos e facilitaram muito a criação dos mapas.

O mapeamento genético com marcadores moleculares é feito essencialmente da mesma forma que o mapeamento feito com os marcadores fenotípicos tradicionais: a cosseparação de dois ou mais marcadores é estudada e as distâncias do mapa se baseiam nas relações de recombinação entre os marcadores. Esses métodos e seu uso no mapeamento são apresentados com mais detalhes nos Capítulos 19 e 20.

Genes podem ser localizados por estudos de associação genômica ampla

A abordagem tradicional para mapear os genes que aprendemos neste capítulo, é examinar os fenótipos dos descendentes nos cruzamentos genéticos ou entre indivíduos em um heredograma, procurando por associações entre a herança de um fenótipo em particular e a herança dos alelos em outros *loci*. Esse tipo de mapeamento é chamado de **análise de ligação**, porque se baseia na detecção da ligação física entre os genes, como medida pela relação de recombinação nos descendentes de um cruzamento. A análise de ligação é uma ferramenta poderosa na análise genética de muitos tipos diferentes de organismos, incluindo a mosca-da-fruta, o milho, os camundongos e os seres humanos.

Outra abordagem alternativa para mapear os genes é realizar **estudos de associação genômica ampla**, procurando por associações não aleatórias entre a presença de um traço e alelos em muitos *loci* diferentes dispersos no genoma. Diferentemente da análise de ligação, essa abordagem não traça a herança dos marcadores genéticos e de um traço em um cruzamento genético ou em uma família. Ela procura por associações entre os traços e conjuntos específicos de alelos em uma *população*.

Imagine que estamos interessados em encontrar os genes que contribuem para o transtorno bipolar, uma doença psiquiátrica caracterizada por grave depressão e mania. Quando uma mutação que predispõe uma pessoa ao transtorno bipolar surge primeiro em uma população, ela ocorre em um cromossomo em particular e está associada a um conjunto específico de alelos naquele cromossomo. No exemplo ilustrado na **Figura 7.18**, a mutação D^- surge primeiro em um cromossomo que tem os alelos A^2, B^2 e C^4, portanto, a mutação *D* está inicialmente ligada aos alelos A^2, B^2 e C^4. Um conjunto específico de alelos ligados como este é chamado de **haplótipo**, e a associação não aleatória entre os alelos em um haplótipo é chamada de **desequilíbrio de ligação**. Como a ligação física entre a mutação bipolar e os outros alelos do haplótipo, o transtorno bipolar e o haplótipo tendem a ser herdados juntos. O *crossing over*, entretanto, rompe a ligação entre os alelos do haplótipo (ver Figura 7.18), reduzindo o desequilíbrio de ligação entre eles. Quanto tempo o desequilíbrio de ligação persiste durante o tempo evolucionário depende da quantidade de recombinação entre os alelos em diferentes *loci*. Quando os *loci* estão muito distantes, o desequilíbrio de ligação se rompe rapidamente, quando os *loci* estão próximos, o *crossing over* é menos comum e o desequilíbrio persiste por mais tempo. O ponto importante é que o desequilíbrio de ligação – a associação não aleatória entre os alelos – fornece informações sobre a distância entre os genes. Uma forte associação entre um traço como o transtorno bipolar e um conjunto de marcadores genéticos ligados indica que é provável que um ou mais genes que contribuem para o transtorno bipolar estejam próximos dos marcadores genéticos.

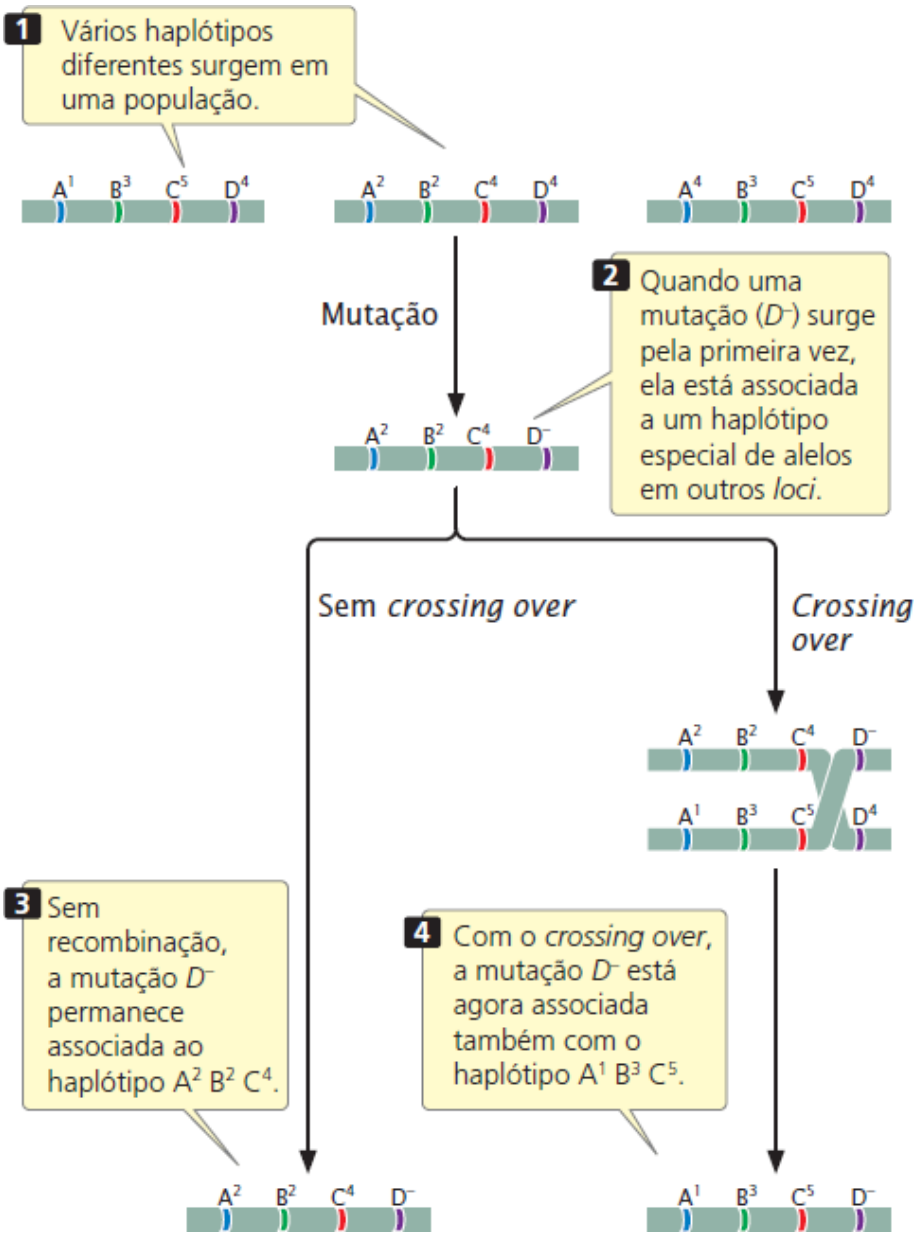


Figura 7.18 Os estudos de associação genômica ampla se baseiam na associação não aleatória de uma mutação (D^-) que produz um traço e genes intimamente ligados que constituem um haplótipo.

Nos últimos anos, os geneticistas mapearam milhões de variantes genéticas chamadas de **polimorfismos de único nucleotídio** (SNPs), que são posições no genoma onde as pessoas variam em uma única base nucleotídica (ver *Capítulo 20*). Lembre-se de que os SNPs foram usados em uma análise de ligação que localizou o gene responsável pela calvície de padrão masculino, discutida na introdução deste capítulo. Agora, é possível estudar de forma rápida e barata o genótipo de pessoas para centenas de milhares ou milhões de SNPs. Esta genotipagem fornece marcadores genéticos necessários para realizar estudos de associação genômica ampla, nas quais os haplótipos de SNPs de pessoas com uma doença em particular, como o transtorno bipolar, são comparados com os haplótipos das pessoas saudáveis. As associações não aleatórias entre SNPs e a doença sugerem que um ou mais genes que contribuem para a doença estão intimamente ligados aos SNPs. Os estudos de associação genômica ampla não localizam genes específicos; eles associam a herança de um traço ou doença com uma região cromossômica específica. Após tal associação ser estabelecida, os geneticistas podem examinar a região cromossômica para os genes que podem ser responsáveis pelo traço. Os estudos de associação genômica ampla são úteis para a descoberta de genes ou regiões cromossômicas que afetam várias doenças genéticas e importantes traços humanos, incluindo o transtorno bipolar, altura, pigmentação da pele, cor dos olhos, peso corporal, coronariopatia, concentrações de lipídios no sangue, diabetes, infartos, densidade óssea e glaucoma, entre outras.

Conceitos

O desenvolvimento de técnicas moleculares para examinar a variação nas sequências de DNA fornece vários marcadores genéticos, que podem ser usados para criar mapas genéticos e estudar relações de ligação. Os estudos de associação genômica ampla examinam a associação não aleatória de marcadores genéticos e fenótipos para localizar os genes que contribuem para a expressão dos traços.

7.4 Os métodos de mapeamento físico são usados para determinar as posições físicas dos genes em cromossomos específicos

Os mapas genéticos revelam as posições relativas dos genes em um cromossomo com base nas frequências de recombinação, mas eles não fornecem informações que permitem que agrupemos os genes ligados em cromossomos particulares. Além disso, as unidades de um mapa genético nem sempre correspondem de forma precisa às distâncias físicas no cromossomo, porque vários fatores além das distâncias físicas entre os genes (como o tipo e sexo do organismo) podem influenciar a recombinação. Por causa dessas limitações, foram desenvolvidos métodos de mapeamento físico que não dependem das frequências de recombinação.

Hibridização de células somáticas

Um método usado para posicionar os genes nos cromossomos é a **hibridização de células somáticas**, que requer a fusão de diferentes tipos de células. A maioria das células somáticas maduras (não sexuais) conseguem realizar apenas um número limitado de divisões e, portanto, não pode crescer de forma contínua. Entretanto, as células que foram alteradas por vírus ou derivadas de tumores e perderam as restrições normais na divisão celular se dividem indefinidamente, podendo ser cultivadas em laboratório para produzir uma **linhagem celular**.

As células de dois tipos de linhagens celulares diferentes podem ser fundidas ao tratá-las com polietilenoglicol ou outros agentes que alteram suas membranas plasmáticas. Após a fusão, a célula tem dois núcleos e é chamada de **heterocárion**. Os núcleos de um heterocárion também se fundem, gerando uma célula híbrida com cromossomos de ambas as linhagens celulares. Se as células humanas e dos camundongos foram misturadas na presença de polietilenoglicol, a fusão gera células híbridas somáticas (**Figura 7.19**). As células híbridas tendem a perder os cromossomos à medida que se dividem e por motivos que não compreendemos, os cromossomos de uma das espécies são preferencialmente perdidos. Nos híbridos de células somáticas de seres humanos e camundongos, os cromossomos humanos tendem a ser perdidos e os cromossomos dos camundongos são retidos. Finalmente, o número de cromossomos estabiliza quando alguns dos cromossomos humanos forem perdidos. A perda de cromossomo é aleatória e difere entre as linhagens celulares. A presença de cromossomos humanos “extra” no genoma do camundongo possibilita avaliar os genes humanos em cromossomos específicos.

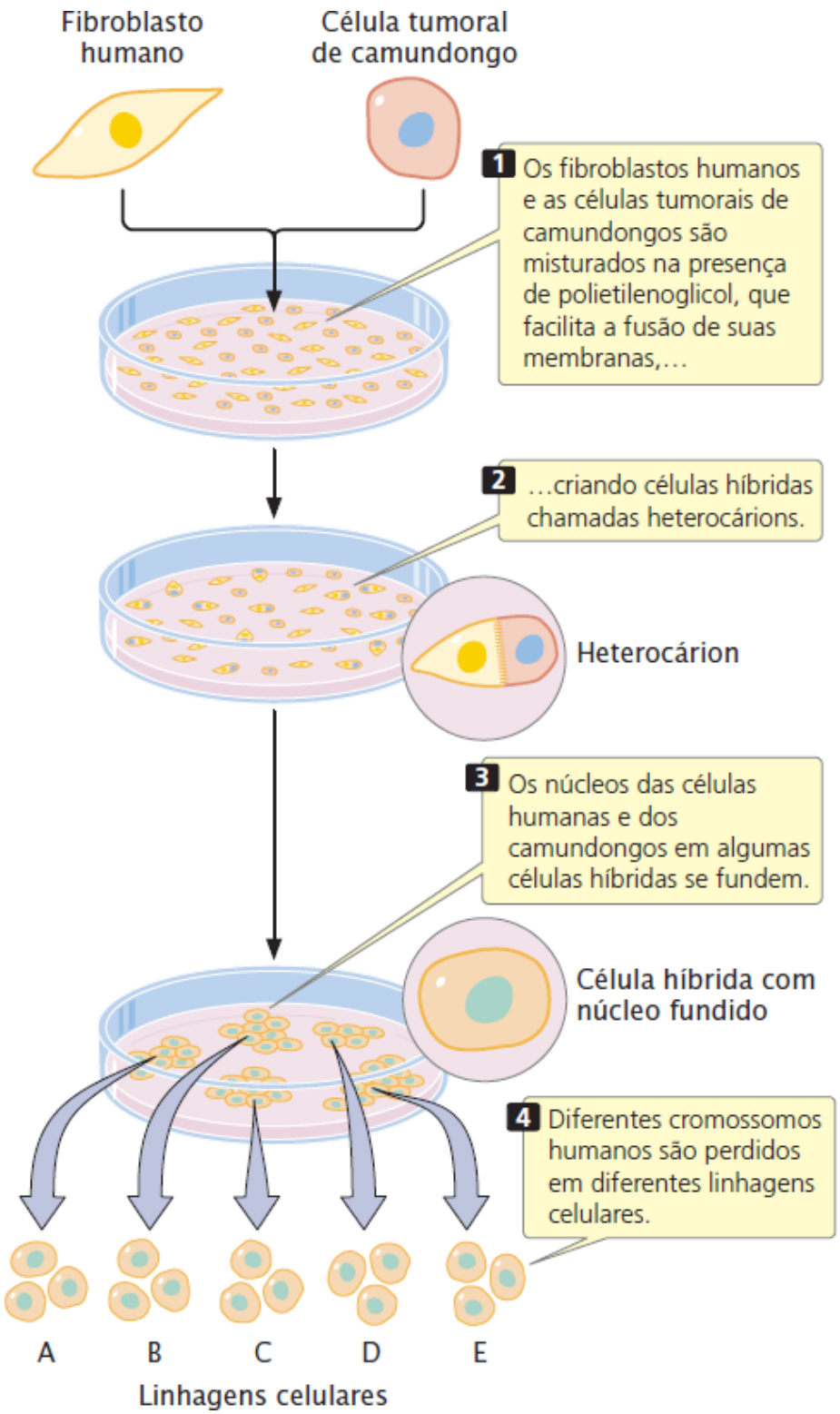


Figura 7.19 A hibridização de células somáticas pode ser usada para determinar qual cromossomo tem um gene de interesse.

Para mapear genes com a hibridização de células somáticas é necessário um painel de diferentes linhagens de células híbridas. Cada linhagem celular é examinada sob microscópio e os cromossomos humanos são identificados. As linhagens celulares do painel são escolhidas de modo que sejam diferentes por conta dos cromossomos humanos que retiveram. Por exemplo, uma linhagem celular poderia ter os cromossomos humanos 2, 4, 7 e 8, enquanto outra poderia ter os cromossomos 4, 19 e 20. Cada linhagem celular no painel é examinada quanto a indícios de um gene humano específico. O gene humano pode ser detectado ao observar o gene propriamente dito (discutido no *Capítulo 19*) ou a proteína que ele produz. A correlação da presença do gene com a existência de cromossomos humanos específicos possibilita, com frequência, que o gene seja designado para o cromossomo certo. Por exemplo, se um gene for detectado em ambas as linhagens celulares mencionadas, ele necessariamente está no cromossomo 4 porque ele é o único cromossomo humano comum às duas linhagens (**Figura 7.20**).

Em alguns casos, a hibridização de células somáticas pode ser usada para posicionar um gene em uma parte específica de um cromossomo. Algumas linhagens de células híbridas carreiam um cromossomo humano com uma mutação como uma deleção ou uma translocação. Se o gene ocorrer em uma linhagem celular com o cromossomo intacto, mas não existir em uma linhagem com

uma deleção cromossômica (uma mutação na qual uma parte de um cromossomo está ausente), o gene deve estar localizado na região deletada (**Figura 7.21**). Da mesma forma, se um gene não existir em um cromossomo, mas aparecer de forma consistente sempre que houver uma translocação (um pedaço de outro cromossomo que foi rompido e preso ao cromossomo em questão), ele obrigatoriamente está na parte translocada do cromossomo.

Cromossomos humanos presentes

Linhagem celular	Produto do gene presente	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X
A	+		+		+			+	+															
B	+	+	+		+				+	+	+	+	+	+										
C	-															+		+		+				
D	+		+		+		+	+	+															
E	-												+								+			
F	+				+															+	+			

Figura 7.20 A hibridização de células somáticas é usada para designar um gene para um cromossomo humano específico. Um painel de seis linhagens celulares, cada uma com um diferente subconjunto de cromossomos humanos, é examinado à procura do produto do gene (como uma enzima). Quatro dessas linhagens (A, B, C e F) têm o produto do gene. O único cromossomo em comum para todas as quatro linhagens celulares é o cromossomo 4, o que indica que o gene está localizado neste cromossomo.

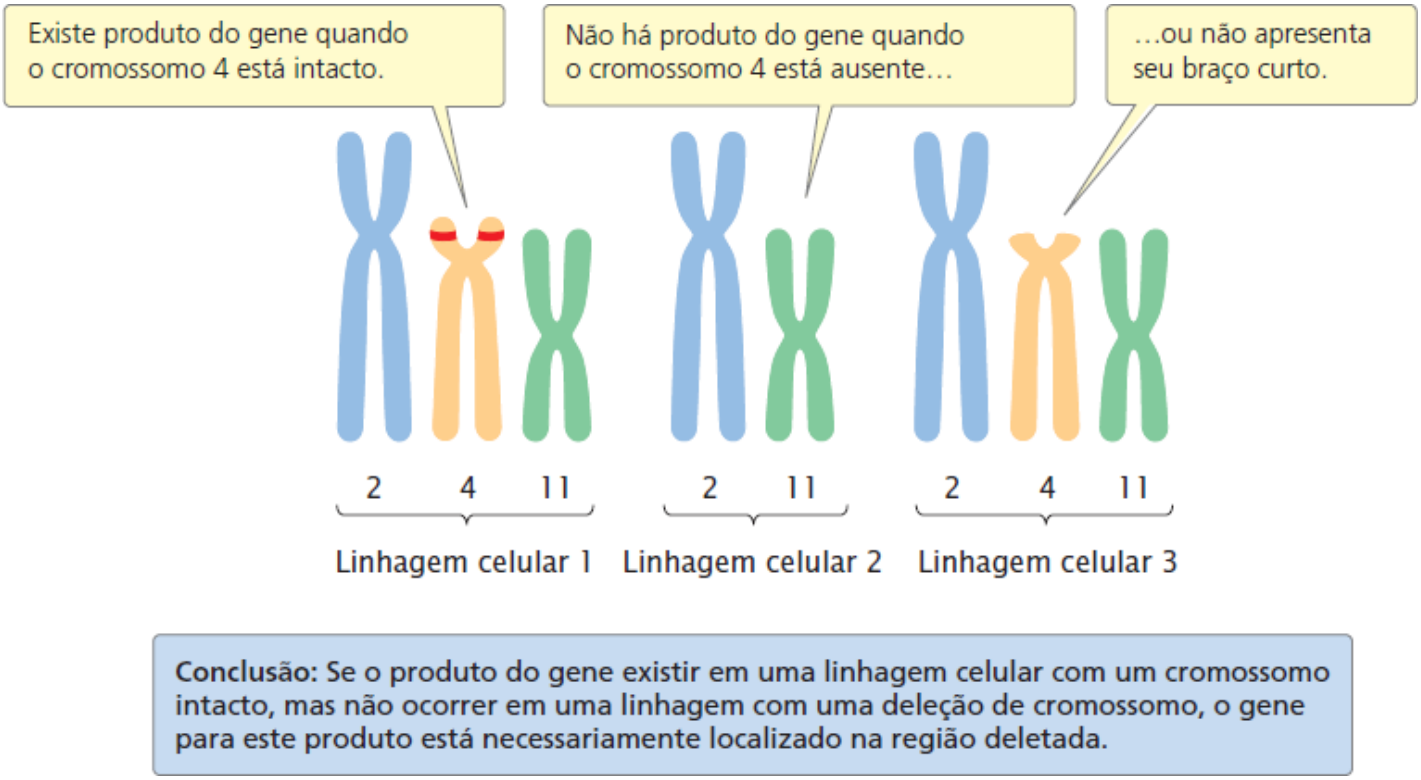


Figura 7.21 Os genes podem ser localizados em uma parte específica de um cromossomo por meio de hibridização de células somáticas.

Problema desenvolvido

Foi criado um painel de linhagens celulares a partir de fusões de células somáticas de seres humanos e camundongos. Foi pesquisada a presença de cromossomos de humanos e a produção de haptoglobina humana (uma proteína) em cada linhagem. Foram obtidos os seguintes resultados:

Linhagem celular	Haptoglobina humana	Cromossomos humanos						
		1	2	3	14	15	16	21

A	–	+	–	+	–	+	–	–
B	+	+	–	+	–	–	+	–
C	+	+	–	–	–	+	+	–
D	–	+	+	–	–	+	–	–

Com base nestes resultados, qual cromossomo humano carreia o gene para haptoglobina?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

O cromossomo carreador do gene para haptoglobina.

Quais informações são fornecidas para solucionar o problema?

- Os cromossomos existentes em cada linhagem celular (a partir do quadro)
- As linhagens celulares que expressam a haptoglobina humana (a partir do quadro).

Etapas para a solução

Primeiro, identifique as linhagens celulares que são positivas para a proteína (haptoglobina humana) e determine os cromossomos que elas têm em comum. As linhagens B e C produzem haptoglobina humana; os únicos cromossomos que elas têm em comum são os cromossomos 1 e 16. A seguir, determine todas as linhagens celulares que têm os cromossomos 1 e 16 e determine se elas produzem haptoglobina. O cromossomo 1 é encontrado nas linhagens celulares A, B, C e D. Se o gene para haptoglobina humana é encontrado no cromossomo 1, esta proteína estaria presente em todas as linhagens. Entretanto, as linhagens A e D não produzem a haptoglobina humana então o gene não pode estar no cromossomo 1. O cromossomo 16 é encontrado apenas nas linhagens B e C, e apenas estas produzem a haptoglobina humana; o gene para esta proteína está no cromossomo 16.

Para praticar mais sobre as hibridizações de células somáticas, resolva o **Problema 37** no final deste capítulo.

Mapeamento de deleção

Outro método para determinar a localização cromossômica de um gene é o **mapeamento de deleção**. Foram desenvolvidos métodos de coloração especiais que revelam os padrões de bandas características nos cromossomos (ver *Capítulo 9*). A ausência de uma ou mais das bandas que estão normalmente em um cromossomo revela a deleção cromossômica. Os genes podem ser designados para regiões dos cromossomos ao estudar a associação entre um fenótipo ou produto do gene e deleções cromossômicas específicas.

No mapeamento da deleção, um indivíduo homozigoto para uma mutação recessiva no gene de interesse é cruzado com um indivíduo heterozigoto para uma deleção (**Figura 7.22**). Se o gene de interesse está na região do cromossomo representado pela deleção (a parte vermelha dos cromossomos na Figura 7.22), então aproximadamente metade dos descendentes apresentará o fenótipo mutante (ver Figura 7.22 A). Se o gene não estiver na região de deleção, então todos os descendentes serão do tipo selvagem (ver Figura 7.22 B).

O mapeamento de deleção é usado para revelar as localizações cromossômicas de vários genes humanos. Por exemplo, a distrofia muscular de Duchenne é uma doença que causa enfraquecimento progressivo e degeneração dos músculos. A partir de seu padrão de herança ligada ao X, sabemos que o alelo mutante responsável por esse distúrbio está no cromossomo X, mas sua localização exata não era certa. O exame de vários pacientes com a doença, que também tinham pequenas deleções, permitiu que os pesquisadores posicionassem o gene em um pequeno segmento do braço curto do cromossomo X. **Resolva o Problema 39**

Mapeamento cromossômico físico por meio de análise molecular

Até o momento, exploramos os métodos que determinam indiretamente a localização cromossômica de um gene ao procurar por seus produtos ou por meio do mapeamento de deleção. Agora os pesquisadores têm as informações e tecnologia para realmente ver onde o gene está. A hibridização *in situ*, descrita com mais detalhes no Capítulo 19, é um método para determinar a localização cromossômica de um gene particular por meio da análise molecular. Este método requer a criação de uma sonda para o gene, que é uma fita única de DNA complementar ao gene de interesse. A sonda é radioativa ou fluorescente sob luz ultravioleta de modo que pode ser visualizada. A sonda se liga à sequência de DNA do gene no cromossomo. A radioatividade ou fluorescência a partir da sonda presa revela a localização do gene em um cromossomo particular.

Além de possibilitar que vejamos onde um gene está localizado em um cromossomo, as técnicas laboratoriais atuais permitem que os pesquisadores identifiquem a localização precisa de um gene em nível de nucleotídio. Por exemplo, com o sequenciamento de

DNA (descrito com mais detalhes no *Capítulo 19*), as distâncias físicas entre os genes podem ser determinadas em números de pares de bases.

Conceitos

Os métodos de mapeamento físico determinam as localizações físicas dos genes nos cromossomos e incluem o mapeamento de deleção, a hibridização de células somáticas, hibridização *in situ* e sequenciamento direto do DNA.

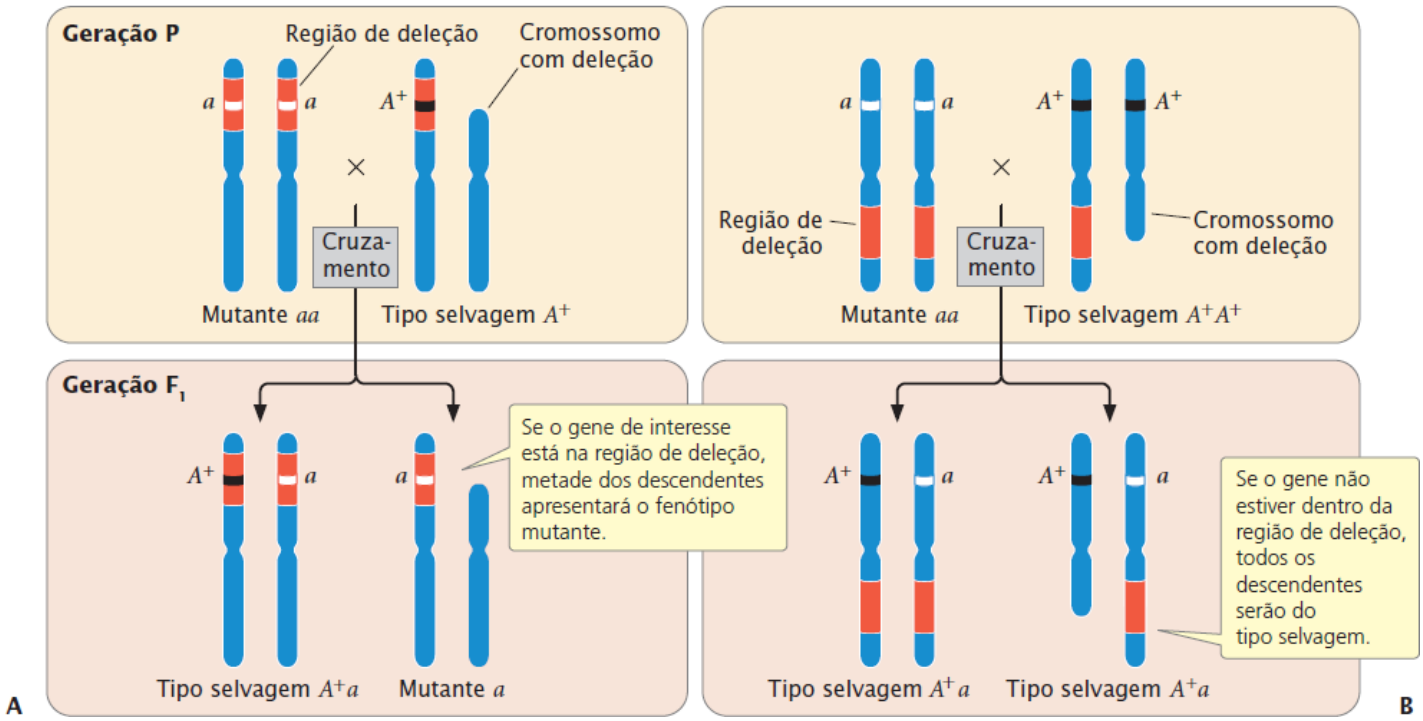


Figura 7.22 O mapeamento da deleção pode ser usado para determinar a localização cromossômica de um gene. Um indivíduo homozigoto para uma mutação recessiva no gene de interesse (*aa*) é cruzado com um indivíduo heterozigoto para uma deleção.

7.5 As taxas de recombinação exibem grande variação

Nos últimos anos, os geneticistas estudaram a variação nas taxas de recombinação e descobriram que seus níveis variam muito – entre espécies, entre e ao longo de cromossomos de uma única espécie, e até entre machos e fêmeas da mesma espécie. Por exemplo, o dobro de recombinação ocorre nos seres humanos, em relação a camundongos e ratos. Dentro do genoma humano, a recombinação varia entre os cromossomos, com os cromossomos 21 e 22 tendo as taxas mais elevadas e os cromossomos 2 e 4 tendo as taxas mais baixas. Os pesquisadores também detectaram diferenças entre os homens e as mulheres nas taxas de recombinação: os cromossomos autossômicos das mulheres sofrem cerca de 50% mais recombinação que os cromossomos autossômicos dos homens.

Os geneticistas encontraram numerosas regiões de *intensa* recombinação (*hotspots*), onde a recombinação é pelo menos 10 vezes acima da média de qualquer outro local do genoma. O genoma humano pode ter cerca de 25.000 a 50.000 regiões de intensa recombinação e aproximadamente 60% de todos os *crossing overs* ocorrem nessa região. Para os seres humanos, as regiões de intensa recombinação (*hotspots*) são encontradas próximas, mas não dentro dos genes ativos. Elas também são detectadas nos genomas de outros organismos. Outras regiões cromossômicas, como as próximas dos centrômeros, apresentam taxas de recombinação reduzidas.

Conceitos

As taxas de recombinação variam entre as espécies, entre e ao longo dos cromossomos e até entre machos e fêmeas.

Resumo dos conceitos

- Os genes ligados não se separam de forma independente em um cruzamento-teste para dois genes completamente ligados (sem *crossing over*), produzindo apenas descendentes não recombinantes. Quando dois genes se separam de forma independente, os descendentes recombinantes e descendentes não recombinantes são produzidos em proporções iguais. Quando dois genes são

ligados e há algum *crossing over* entre eles, são produzidos mais descendentes não recombinantes que recombinantes

- A frequência de recombinação é calculada ao somar o número de descendentes recombinantes, dividir pelo total de descendentes produzidos no cruzamento e multiplicar por 100%. Ela é a metade da frequência do *crossing over* e a frequência máxima de gametas recombinantes é de 50%
- Acoplamento e repulsão se referem à disposição dos alelos em um cromossomo. O fato de os genes estarem em acoplamento ou repulsão determina que combinação de fenótipos será mais frequente nos descendentes de um cruzamento-teste
- A recombinação intercromossômica ocorre entre genes localizados em diferentes cromossomos por meio da separação aleatória dos cromossomos na meiose. A recombinação intracromossômica ocorre entre genes localizados no mesmo cromossomo pelo *crossing over*
- Um teste qui-quadrado de independência pode ser usado para determinar se os genes estão ligados
- As taxas de recombinação podem ser usadas para determinar a ordem relativa de genes e a distância entre eles em um cromossomo. Uma porcentagem de recombinação é igual a uma unidade de mapa. Os mapas que se baseiam nas taxas de recombinação são chamados de mapas genéticos; os mapas que se baseiam nas distâncias físicas são chamados de mapas físicos
- Os mapas genéticos podem ser construídos ao examinar as taxas de recombinação a partir de vários cruzamentos de dois pontos ou ao examinar os descendentes de um cruzamento-teste de três pontos
- Alguns *crossing overs* múltiplos passam despercebidos, assim, os mapas genéticos que usam as taxas de recombinação subestimam as verdadeiras distâncias físicas entre os genes
- Os genes humanos podem ser mapeados ao examinar a cosseparação dos traços nas linhagens
- Um *lod score* é o logaritmo da razão da probabilidade de obter os descendentes observados com a hipótese de ligação, para a probabilidade de obter os descendentes observados com a hipótese de segregação independente. Um *lod score* de 3 ou maior em geral é considerado um indício convincente de ligação
- As técnicas moleculares que permitem a detecção de diferenças variáveis na sequência de DNA têm facilitado muito o mapeamento genético
- Os estudos de associação genômica ampla localizam genes que afetam traços particulares ao examinar a associação não aleatória de um traço com os marcadores genéticos por meio do genoma
- O sequenciamento nucleotídico é outro método de mapeamento físico dos genes
- As taxas de recombinação variam muito, diferem entre espécies, entre e ao longo de cromossomos dentro de uma única espécie e até entre machos e fêmeas da mesma espécie.

Termos importantes

Análise de ligação
CentiMorgan (cM)
Coeficiente de coincidência
Configuração *cis* (acoplamento)
Configuração *trans* (repulsão)
Cruzamento-teste de dois pontos
Cruzamento-teste de três pontos
Descendentes não recombinantes (parentais)
Descendentes recombinantes
Desequilíbrio de ligação
Estudos de associação genômica ampla
Frequência de recombinação
Função de mapeamento
Gameta não recombinante (parental)
Gameta recombinante
Genes ligados
Grupo de ligação
Haplótipo
Heterocárion
Hibridização de células somáticas
Interferência
Linhagem celular
Lod score (logaritmo de probabilidade)
Mapa físico
Mapa genético
Mapeamento de deleção
Marcador genético

Polimorfismo de único nucleotídio (SNP)
Unidade de mapa (m.u.)

1. c.
2. Repulsão.
3. Os mapas genéticos se baseiam nas taxas de recombinação, enquanto os mapas físicos se baseiam nas distâncias físicas.
4. $\frac{m^+\ p^+\ s^+}{m\ \ p\ \ s}\ \frac{m^+\ p\ s}{m\ \ p\ s}\ \frac{m\ \ p^+\ s^+}{m\ \ p\ \ s}\ \frac{m^+\ p^+\ s}{m\ \ p\ \ s}\ \frac{m\ p\ s^+}{m\ p\ s}\ \frac{m^+\ p\ s^+}{m\ \ p\ s}\ \frac{m\ p^+\ s}{m\ p\ \ s}\ \frac{m\ p\ s}{m\ p\ s}$
5. O *locus c*.
6. b.

Problemas desenvolvidos

Problema 1

Nos porquinhos-da-índia, o pelo branco (*w*) é recessivo para o pelo preto (*W*) e o fio ondulado (*v*) é recessivo para o fio liso (*V*). Um criador cruza um porquinho-da-índia homozigoto para pelo branco e fio ondulado com um porquinho-da-índia com pelo preto e fio liso. Os F_1 são, então, cruzados com porquinhos com pelo branco e fio ondulado em uma série de cruzamentos-teste. Foram produzidos os seguintes descendentes a partir destes cruzamentos-teste:

Preto, liso	30
Preto, ondulado	10
Branco, liso	12
Branco, ondulado	31
Total	83

a. Os genes que determinam cor e tipo do pelo se separam de forma independente? Faça testes qui-quadrado para verificar sua hipótese.
b. Se os genes não se separam de forma independente, qual é a frequência de recombinação entre eles?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

a. Se os genes se separam de forma independente, junto com o valor de qui-quadrado, *g.l.* e valor de P para avaliar sua hipótese.
b. Se os genes não se separam de forma independente, a frequência de recombinação entre eles.

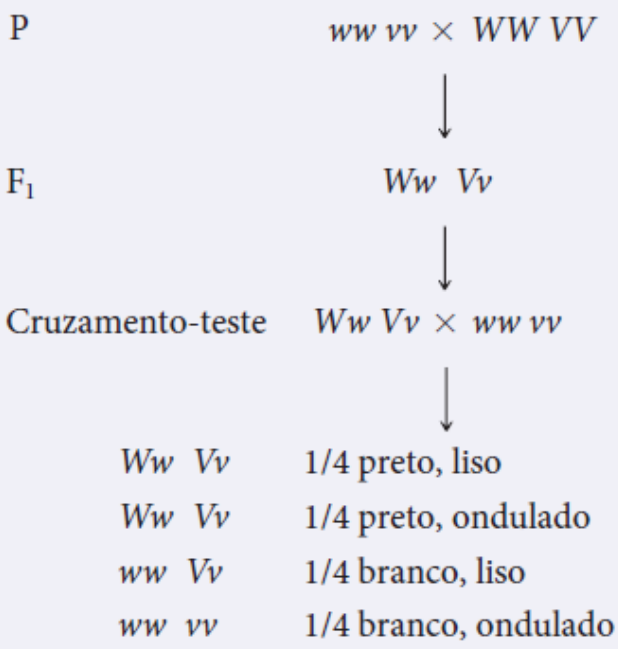
Quais informações são fornecidas para solucionar o problema?

- O pelo branco (*w*) é recessivo para o pelo preto (*W*) e o fio ondulado (*b*) é recessivo para o fio liso (*V*)
- Os números dos diferentes tipos de descendentes de uma série de cruzamentos-teste.

Para a solução deste problema, revise:
Crossing over com genes ligados, *Como calcular a frequência de recombinação* e *Análise de segregação independente*, na Seção 7.2.

Etapas para a solução

a. Considerando que há segregação independente, esboce os cruzamentos feitos pelo criador:



Lembrete: Com a segregação independente, esperamos números iguais de descendentes não recombinantes e recombinantes:

Como foram produzidos 83 descendentes nos cruzamentos-teste, esperamos 1/4 × 83 = 20,75 de cada. Os números observados dos descendentes do cruzamento-teste (30, 10, 12, 31) não parecem se ajustar aos números esperados (20,75; 20,75; 20,75; 20,75), então pode não ter ocorrido segregação independente.

Dica: Veja Figura 7.9 para revisar como fazer um teste qui-quadrado de independência.

Para testar a hipótese, faça um teste qui-quadrado de independência. Construa um quadro com os genótipos do primeiro *locus* na linha superior e os genótipos do segundo *locus* na coluna à esquerda. Calcule os valores totais para as linhas e colunas e o total final.

	Ww	*ww*	Total das linhas
Vv	30	12	42
vv	10	31	41
Total das colunas	40	43	83 ← Total final

O valor esperado para cada célula do quadro é calculado com a fórmula a seguir:

$$\text{Número esperado} = \frac{\text{Total de linhas} \times \text{Total de colunas}}{\text{Total final}}$$

Usando esta fórmula, encontramos os valores esperados (indicado entre parênteses):

	Ww	*ww*	Total das linhas
Vv	30 (20,24)	12 (21,76)	42
vv	10 (19,76)	31 (21,24)	41
Total das colunas	40	43	83 ⟵ Total final

Usando estes valores observados e esperados, encontramos o valor de qui-quadrado:

$$\chi^2 = \sum \frac{(\text{Observados} - \text{Esperados})^2}{\text{Esperados}}$$

$$= \frac{(30 - 20{,}24)^2}{20{,}24} + \frac{(10 - 19{,}76)^2}{19{,}76} + \frac{(12 - 21{,}76)^2}{21{,}76} + \frac{(31 - 21{,}24)^2}{21{,}24}$$

$$= 4{,}71 + 4{,}82 + 4{,}38 + 4{,}48 = 18{,}39$$

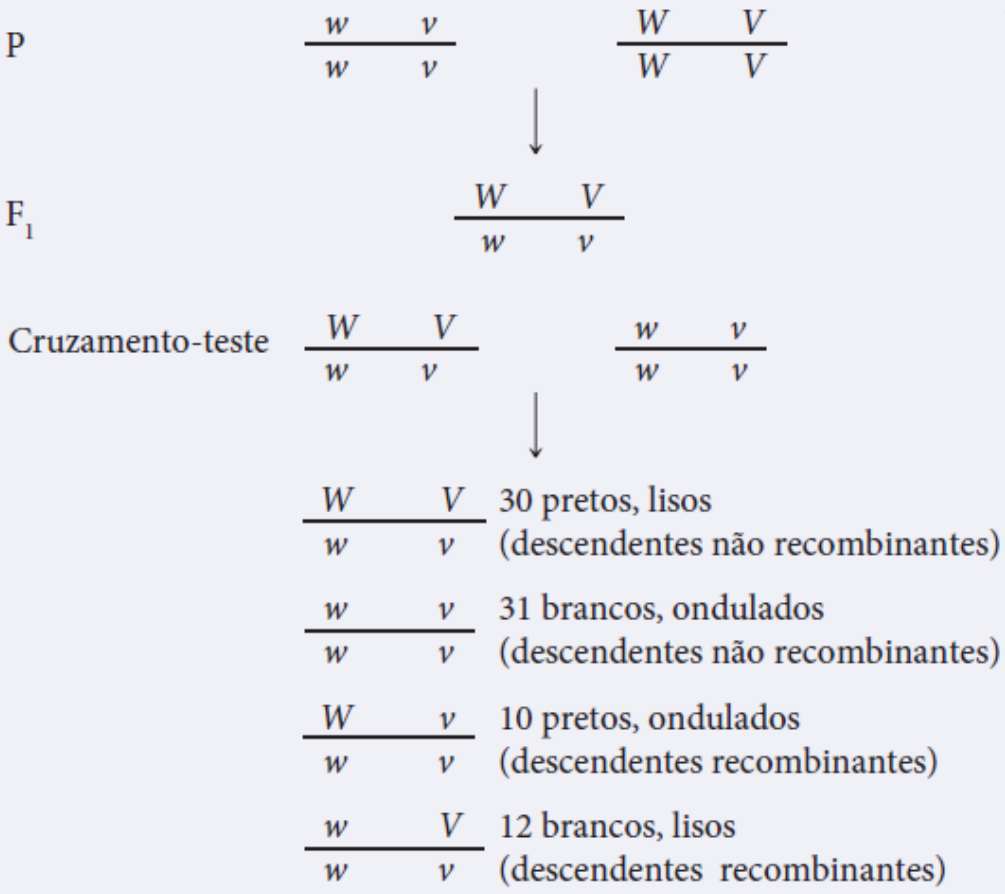
Os graus de liberdade do teste qui-quadrado de independência são *gl* = (número de linhas – 1) × (número de colunas – 1). Existem duas linhas e colunas, então os graus de liberdade são:

$$gl = (2 - 1) \times (2 - 1) = 1 \times 1 = 1$$

No Quadro 3.5, a probabilidade associada a um valor de qui-quadrado de 18,39 e 1 grau de liberdade é menor que 0,005, indicando que é muito improvável que o acaso seja responsável pelas diferenças entre os números observados e os números esperados com a segregação independente. Os genes para cor e tipo de pelo não se separam de forma independente.

Lembrete: A probabilidade associada ao valor de qui-quadrado é a probabilidade de que o acaso seja responsável pela diferença entre os valores observados e esperados.

b. Para determinar as frequências de recombinação, identifique os descendentes recombinantes. Usando a notação para os genes ligados, escreva os cruzamentos:



A frequência de recombinação é:

Frequência de recombinação =

$$\frac{10 + 12}{30 + 31 + 10 + 12} \times 100\% = \frac{22}{83} \times 100\% = 26{,}5$$

Lembrete: A frequência de recombinação = $\frac{\text{número de descendentes recombinantes}}{\text{número total de descendentes}} \times 100\%$.

Problema 2

Alguns cruzamentos de dois pontos foram feitos entre sete *loci* (*a*, *b*, *c*, *d*, *e*, *f* e *g*), produzindo as frequências de recombinação a seguir. Usando-as, faça o mapa dos sete *loci*, mostrando seus grupos de ligação, a ordem dos *loci* em cada grupo de ligação e as distâncias entre os *loci* de cada grupo:

Loci	**Frequência de recombinação (%)**	***Loci***	**Frequência de recombinação (%)**
a e *b*	10	*c* e *d*	50
a e *c*	50	*c* e *e*	8
a e *d*	14	*c* e *f*	50
a e *e*	50	*c* e *g*	12
a e *f*	50	*d* e e	50
a e *g*	50	*d* e *f*	50

b e *c*	50	*d* e *g*	50
b e *d*	4	*e* e *f*	50
b e *e*	50	*e* e *g*	18
b e *f*	50	*f* e *g*	50
b e *g*	50		

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Os grupos de ligação para os sete *loci*, a ordem dos *loci* em cada grupo de ligação e as distâncias de mapa entre os *loci*.

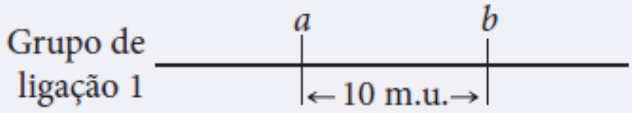
Quais informações são fornecidas para solucionar o problema?
As frequências de recombinação para cada par de *loci*.

Para a solução deste problema, revise:
Como construir um mapa genético com o uso de cruzamentos-teste de dois pontos, na Seção 7.2.

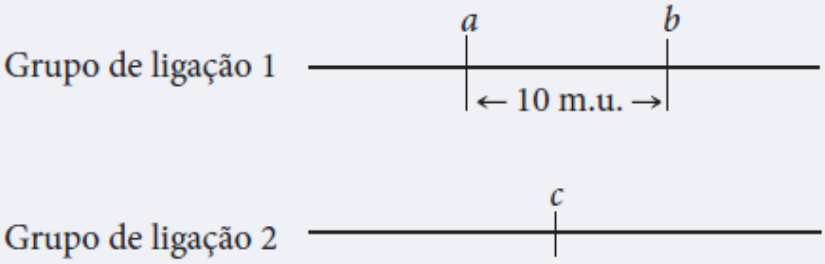
Etapas para a solução

Para resolver este problema, lembre-se de que a recombinação de 1% é igual a 1 unidade de mapa. A frequência de recombinação entre *a* e *b* é de 10%, então estes *loci* estão no mesmo grupo de ligação, a uma distância aproximada de 10 m.u. (unidades de mapa).

Dica: Uma frequência de recombinação de 50% significa que os genes em dois *loci* estão se separando de forma independente (localizados em diferentes grupos de ligação).

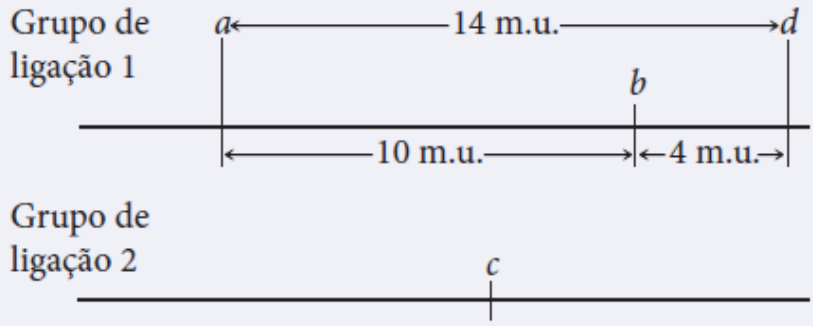


A frequência de recombinação entre *a* e *c* é 50%, indicando que *c* deve estar em um segundo grupo de ligação.

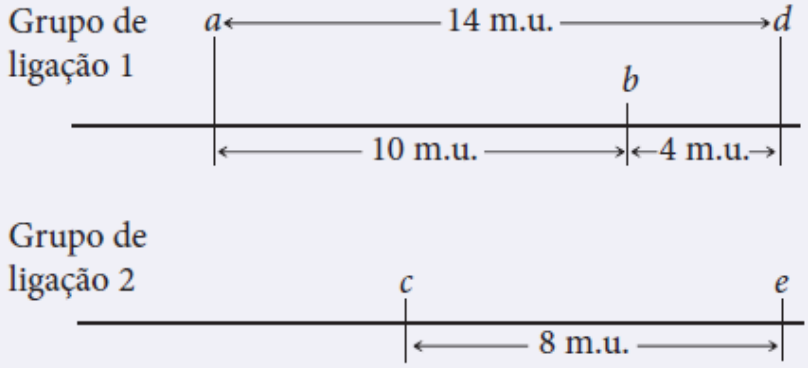


A frequência de recombinação entre *a* e *d* é 14%, então *d* está localizado no grupo de ligação 1. O *locus d* está 14 m.u. à direita ou à esquerda do *locus a*? O *locus d* está 14 m.u. à esquerda de *a*, então a distância entre *b* e *d* deve ser 10 m.u. + 14 m.u. = 24 m.u. Por outro lado, se *d* está à direita de *a*, então a distância entre *b* e *d* deve ser 14 m.u. – 10 m.u. = 4 m.u. A frequência de recombinação *b*-*d* é 4%, então *d* está 14 m.u. à direita de *a*. O mapa atualizado é:

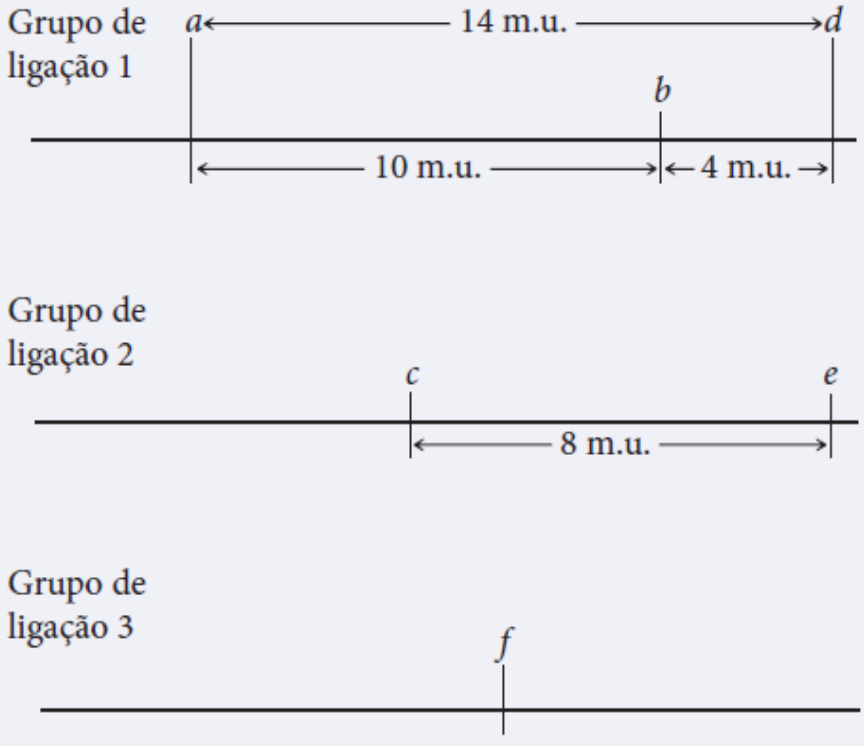
Dica: Para determinar se o *locus d* está à direita ou à esquerda do *locus a*, observe a distância entre *b* e *d*.



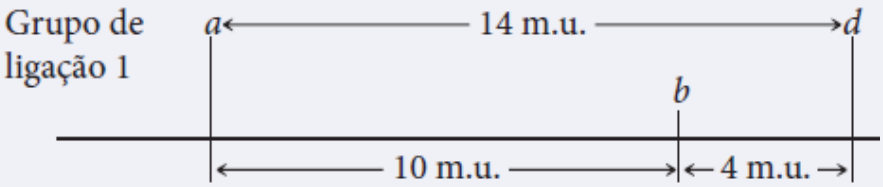
As frequências de recombinação entre cada um dos *loci* *a*, *b* e *d* e *locus* *e* são todas 50%; então *e* não está no grupo de ligação 1 com *a*, *b* e *d*. A frequência de recombinação entre *e* e *c* é 8 m.u., então *e* está no grupo de ligação 2:

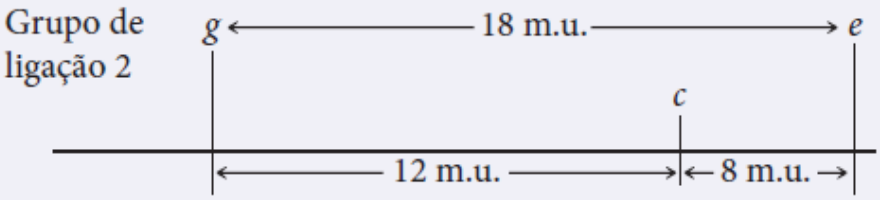


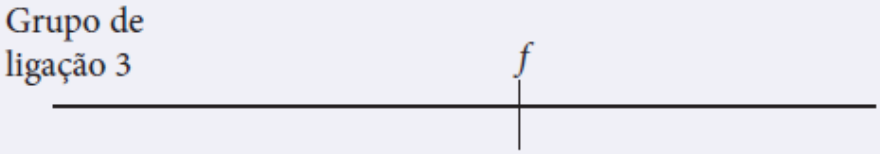
Existe uma recombinação de 50% entre *f* e todos os outros genes; então *f* obrigatoriamente pertence a um terceiro grupo de ligação:



Finalmente, a posição do *locus* *g* com relação aos outros genes. As frequências de recombinação entre *g* e os *loci* *a*, *b* e *d* são todas 50%, então *g* não está localizado no grupo de ligação 1. A frequência de recombinação entre *g* e *c* é 14%, então *g* faz parte do grupo de ligação 2. Para determinar se *g* está a 12 m.u. à direita ou à esquerda de *c*, consulte a frequência de recombinação entre *g*-*e*. Como ela é de 18%, *g* deve estar à esquerda de *c*:





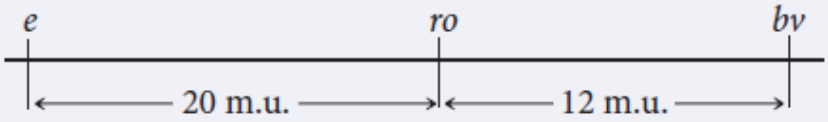


Observe que a distância *g* para *e* (18 m.u.) é menor que a soma das distâncias de *g* para *c* (12 m.u.) e *c* para *e* (8 m.u.) por causa dos *crossing overs* duplos não detectáveis entre *g* e *e*.

Lembrete: Como alguns *crossing overs* duplos podem passar não detectados, a distância entre dois genes (como *g* e *e*) pode ser menor que a soma das distâncias menores (como de *g* para *c* e *c* para *e*).

Problema 3

A cor ébano do corpo (*e*), os olhos rugosos (*ro*) e as cerdas curtas (*bv*) são três mutações recessivas que ocorrem nas moscas-da-fruta. Os *loci* para essas mutações foram mapeados e estão separados pelas seguintes distâncias de mapa:



A interferência entre esses genes é 0,4.

Uma mosca com corpo cor de ébano, olhos rugosos e cerdas duplas é cruzada com uma mosca que é homozigota para os traços do tipo selvagem. As fêmeas F_1 resultantes são usadas para um cruzamento-teste com machos com corpo cor de ébano, olhos rugosos e cerdas duplas, produzindo 1.800 descendentes. Indique os números esperados de fenótipos nos descendentes do cruzamento-teste.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Os números esperados dos diferentes fenótipos produzidos pelo cruzamento-teste.

Quais informações são fornecidas para solucionar o problema?

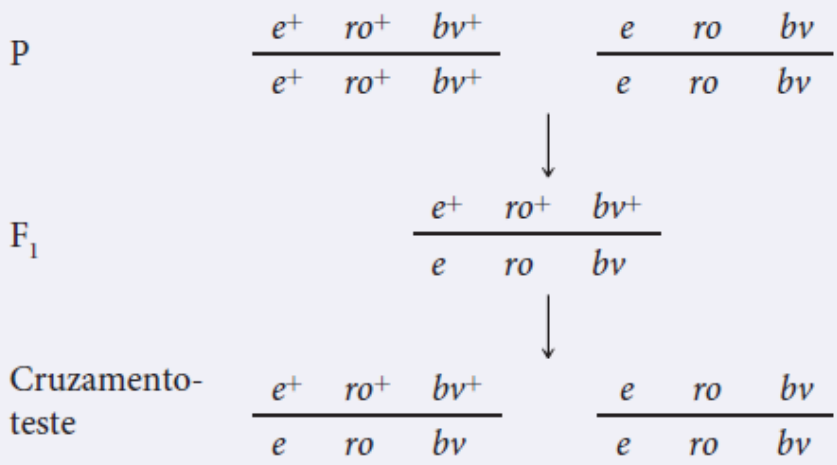
- As distâncias de mapa entre os três *loci*
- A interferência entre os *loci*
- É feito um cruzamento-teste e são produzidos 1.800 descendentes.

Para a solução deste problema, revise:

Como construir um mapa genético com o uso de cruzamento-teste de três pontos e o *Problema desenvolvido*, na Seção 7.3.

Etapas para a solução

Os cruzamentos são:



Neste caso, sabemos que *ro* está no *locus* do meio porque os genes foram mapeados. Serão produzidas oito classes de descendentes a partir desse cruzamento:

Dica: A ordem dos genes é fornecida pelo mapa genético.

e^+		ro^+		bv^+	Não recombinante
e		*ro*		*bv*	Não recombinante
e^+	/	*ro*	*bv*		*Crossing over* único entre *e* e *ro*
e	/	ro^+	bv^+		*Crossing over* único entre *e* e *ro*
e^+		ro^+	/	*bv*	*Crossing over* único entre *ro* e *bv*
e		*ro*	/	bv^+	*Crossing over* único entre *ro* e *bv*
e^+	/	*ro*	/	bv^+	*Crossing over* duplo
e	/	ro^+	/	*bv*	*Crossing over* duplo

Use as distâncias de mapa para determinar os números de cada tipo, começando com os *crossing overs* duplos. O número esperado de *crossing overs* duplos é igual ao produto das probabilidades de *crossing over* único:

Número esperado de *crossing overs* duplos = 0,20 × 0,12 × 1.800 = 43,2

Entretanto, existe alguma interferência, então o número observado de *crossing overs* duplos será menor que o esperado. A interferência é 1 – coeficiente de coincidência, sendo o coeficiente de coincidência:

Lembrete: A existência de interferência significa que nem todos os *crossing overs* duplos esperados serão observados.

Coeficiente de coincidência = 1 – interferência

A interferência é 0,4, então o coeficiente de coincidência é igual a 1 – 0,4 = 0,6. Lembre-se de que o coeficiente de coincidência é:

$$\text{Coeficiente de coincidência} = \frac{\text{Número de } \textit{crossing overs} \text{ duplos observados}}{\text{Número de } \textit{crossing overs} \text{ duplos esperados}}$$

Reorganizando esta equação, temos:

Número de *crossing overs* duplos observados	=	Coeficiente de coincidência	×	Número de *crossing overs* duplos esperados

Número de *crossing overs* duplos observados = 0,6 × 43,2 = 26

Será observado um total de 26 *crossing overs* duplos. Como existem duas classes de *crossing overs* duplos (*e*$^{\pm}$/*ro*/*bv*$^{\pm}$ e *e*/*ro*$^{\pm}$/*bv*), esperamos observar 13 de cada classe.

Dica: Para obter o número de descendentes do *crossing over* único, subtraia o número de descendentes de *crossing over* duplo do número total resultante da recombinação.

A seguir, determinamos o número de descendentes de cruzamento único. O mapa genético indica que a distância entre *e* e *ro* é 20 m.u.; então são esperados 360 descendentes (20% de 1.800) como resultado da recombinação entre esses *loci*, alguns descendentes de *crossing over* único e alguns de *crossing over* duplo. Já determinamos que o número de descendentes do *crossing over* duplo é 26; então o número de descendentes resultantes de um *crossing over* único entre *e* e *ro* é 360 – 26 = 334, que será dividido igualmente entre os dois fenótipos de *crossing over* único (*e*/*ro*$^{\pm}$/*bv*$^{\pm}$ e *e*$^{\pm}$/*robv*).

A distância entre *ro* e *bv* é 12 m.u., então o número de descendentes resultantes da recombinação entre esses dois genes é 0,12 × 1.800 = 216. Novamente, alguns desses recombinantes serão descendentes de *crossing over* único e alguns de *crossing over* duplo. Para determinar o número de descendentes resultantes de um *crossing over* único, subtraia os *crossing overs* duplos: 216 – 26 = 190. Os descendentes de *crossing over* único serão divididos entre os dois fenótipos de *crossing over* único (*e*$^{\pm}$*ro*$^{\pm}$/*bv* e *ero*/*bv*$^{\pm}$), então teremos 190/2 = 95 de cada um desses fenótipos. Os descendentes remanescentes serão não recombinantes, e podem ser obtidos pela subtração: 1.800 – 26 – 334 – 190 = 1.250; existem dois não recombinantes (*e*$^{\pm}$ *ro*$^{\pm}$ *bv*$^{\pm}$ e *e ro bv*); então teremos 1.250/2 = 625 de cada. Os números dos vários fenótipos estão listados aqui:

Dica: Não se esqueça de subtrair os *crossing overs* duplos do número total de recombinantes.

Dica: O número de recombinantes pode ser obtido pela subtração.

e^{+}		*ro*$^{+}$	*bv*$^{+}$	625	Não recombinante
e		*ro*	*bv*	625	Não recombinante
e^{+}	/	*ro*	*bv*	167	*Crossing over*

						único entre *e* e *ro*
e	/	*ro*$^+$		*bv*$^+$	167	*Crossing over* único entre *e* e *ro*
e$^+$		*ro*$^+$	/	*bv*	95	*Crossing over* único entre *ro* e *bv*
e		*ro*	/	*bv*$^+$	95	*Crossing over* único entre *ro* e *bv*
e$^+$	/	*ro*	/	*bv*$^+$	13	*Crossing over* duplo
e	/	*ro*$^+$	/	*bv*	13	*Crossing over* duplo
Total					1.800	

Questões de compreensão

Seção 7.1

1. O que significa o termo recombinante? Quais são as duas causas de recombinação?

Seção 7.2

2. Em um cruzamento-teste para dois genes, quais tipos de gametas são produzidos com (a) ligação completa, (b) segregação independente e (c) ligação incompleta?
3. Qual efeito o *crossing over* tem sobre a ligação?
4. Por que a frequência dos gametas recombinantes é sempre a metade da frequência do *crossing over*?
5. Qual a diferença entre os genes na configuração de acoplamento e os genes na configuração de repulsão? Como a disposição dos genes ligados (se eles estão em acoplamento ou repulsão) afeta os resultados de um cruzamento genético?
6. Como você faria um teste para verificar se dois genes estão ligados?
7. Qual é a diferença entre um mapa genético e um mapa físico?
8. Por que as frequências de recombinação calculadas entre os pares de *loci* que estão distantes subestimam as verdadeiras distâncias genéticas entre os *loci*?

Seção 7.3

9. Explique como determinar, usando os números de descendentes de um cruzamento de três pontos, qual dos três *loci* ligados está no meio.
10. O que a interferência nos diz sobre o efeito de um *crossing over* sobre outro?

Seção 7.4

11. O que é um *lod score* e como ele é calculado?
12. Liste alguns dos métodos para mapeamento físico dos genes e explique como eles são usados para posicionar os genes nos cromossomos.

Questões e problemas aplicados

Introdução

*13. No início deste capítulo, falou-se da pesquisa pelos genes que determinam a alopecia androgenética nos seres humanos. Em 1916, Dorothy Osborn sugeriu que este tipo de calvície é um traço influenciado pelo sexo (ver *Capítulo 5*), dominante nos homens e recessivo nas mulheres. Outras pesquisas sugerem que a alopecia androgenética é um traço recessivo ligado ao X.

Você esperaria ver segregação independente entre os marcadores genéticos no cromossomo X e a alopecia androgenética se (a) este tipo de calvície fosse influenciada pelo sexo e (b) se ela fosse recessiva ligada ao X? Explique sua resposta.

Seção 7.2

14. No caracol *Cepaea nemoralis*, um alelo autossômico responsável pelo casco listrado (B^B) é recessivo para o alelo de casco não listrado (B^O). Os genes em um *locus* diferente determinam a cor de fundo do casco, sendo amarelo (C^Y) recessivo para marrom (C^{Bw}). Um caracol listrado amarelo é cruzado com um caracol não listrado marrom homozigoto. Então, eles são cruzados com caracóis amarelos listrados (cruzamento-teste).
    a. Quais serão os resultados do cruzamento-teste se os *loci* que controlam as listras e a cor estiverem ligados sem *crossing over*?
    b. Quais serão os resultados do cruzamento-teste se os *loci* se separam de forma independente?
    c. Quais serão os resultados do cruzamento-teste se os *loci* estiverem ligados e 20 m.u. distantes um do outro?

*15. No bicho-da-seda (*Bombyx mori*), os traços olhos vermelhos (*re*) e asas brancas listradas (*wb*) são codificados por dois alelos mutantes que são recessivos em relação aos que produzem os traços do tipo selvagem (re^+ e wb^+); esses dois genes estão no mesmo cromossomo. Um bicho-da-seda homozigoto para olhos vermelhos e asas brancas listradas é cruzado com um homozigoto para os traços do tipo selvagem. Os descendentes F_1 têm asas e olhos normais. Eles são cruzados com bichos com olhos vermelhos e asas listradas brancas em um cruzamento-teste. Os descendentes desse cruzamento são:

Olhos do tipo selvagem, asas do tipo selvagem	418
Olhos vermelhos, asas do tipo selvagem	19
Olhos do tipo selvagem, asas brancas listradas	16
Olhos vermelhos, asas brancas listradas	426

    a. Quais proporções fenotípicas seriam esperadas se os genes para olhos vermelhos e asas brancas listradas estivessem localizados em diferentes cromossomos?
    b. Qual a porcentagem de recombinação entre os genes para olhos vermelhos e para asas brancas listradas?

*16. Uma geneticista descobre uma nova mutação na *Drosophila melanogaster* que faz com que as moscas tremam e se sacudam. Ela chama tal mutação de espástica (*sps*), determina que é causada por um gene recessivo autossômico e quer determinar se o gene que codifica o traço espástico está ligado ao gene recessivo para asas vestigiais (*vg*). Uma mosca homozigota para os traços espástico e vestigial é cruzada com uma mosca homozigota para os traços do tipo selvagem. As fêmeas F_1 resultantes são usadas em um cruzamento-teste, obtendo-se as seguintes moscas:

vg^+	sps^+	230
vg	*sps*	224
vg	sps^+	97
vg^+	*sps*	99
Total		650

Os genes responsáveis pelas asas vestigiais e a mutação espástica estão ligados? Faça um teste qui-quadrado de independência para determinar se os genes se separam de forma independente.

*17. Nos pepinos, as folhas em formato de coração (*hl*) são recessivas para as folhas normais (*Hl*), e ter numerosos frutos espinhosos (*ns*) é recessivo para ter poucos frutos espinhosos (*Ns*). Os genes para o formato da folha e para o número de espinhos estão localizados no mesmo cromossomo e os achados para mapear os experimentos indicam que eles estão a uma distância de 32,6 m.u. Um pepino com folhas em formato de coração e numerosos espinhos é cruzado com uma planta homozigota para folhas normais e poucos espinhos. Os descendentes F_1 são cruzados com plantas de folhas em formato de coração e numerosos espinhos. Quais fenótipos e proporções fenotípicas são esperados nos descendentes deste cruzamento?

18. Nos tomates, a planta alta (*D*) é dominante sobre a planta anã (*d*) e a fruta lisa (*P*) é dominante sobre a fruta pubescente (*p*), que é coberta por finos pelos. Um fazendeiro tem duas plantas de tomates altas e fruta lisa, que chamaremos de planta A e planta B. Ele cruza as duas plantas com a mesma planta anã e pubescente e obtém os seguintes números de descendentes:

	Descendentes de	
	Planta A	**Planta B**
Dd Pp	122	2
Dd pp	6	82
dd Pp	4	82
dd pp	124	4

    a. Quais são os genótipos das plantas A e B?

b. Os *loci* que determinam a altura e a pubescência da planta estão ligados? Se sim, qual é a porcentagem de recombinação entre eles?

c. Explique por que diferentes proporções de descendentes são produzidas quando as plantas A e B são cruzadas com a mesma planta pubescente anã.

19. Os alelos *A* e *a* estão em um *locus* no mesmo cromossomo, assim como há um *locus* com os alelos *B* e *b*. *Aa Bb* é cruzado com *aa bb*, e os seguintes descendentes são produzidos:

Aa Bb	5
Aa bb	45
aa Bb	45
aa bb	5

Que conclusão pode ser tirada sobre a disposição dos genes no cromossomo no genitor *Aa Bb*?

ANÁLISE DE DADOS

20. Daniel McDonald e Nancy Peer determinaram que o *eyespot* (um ponto claro no centro do olho) de quatro tenébrios é causado por um gene ligado ao X (*es*) que é recessivo para o alelo para ausência do *eyespot* (es^+). Eles fizeram vários cruzamentos para determinar a distância entre o gene para *eyespot* e um gene dominante ligado ao X para listrado (*St*), que produz listras brancas nas fêmeas e age como um gene recessivo letal (é letal quando homozigoto nas fêmeas e hemizigoto nos machos). Foi feito o cruzamento a seguir. (D. J. McDonald and N. J. Peer. 1961. *Journal of Heredity* 52:261-264.)

$$\text{♀}\ \frac{es^+\quad St}{es\quad St^+} \times \frac{es\quad St^+}{Y}\ \text{♂}$$

$\frac{es^+\quad St}{es\quad St^+}$	1.630
$\frac{es\quad St^+}{es\quad St^+}$	1.665
$\frac{es\quad St}{es\quad St^+}$	935
$\frac{es^+\quad St^+}{es\quad St^+}$	1.005
$\frac{es\quad St^+}{Y}$	1.661
$\frac{es^+\quad St^+}{Y}$	1.024

a. Quais descendentes eram recombinantes e quais eram não recombinantes?

b. Calcule a frequência de recombinação entre *es* e *St*.

c. Existem genótipos potenciais ausentes entre os descendentes do cruzamento? Se esse for o caso, quais são e por quê?

*21. As taxas de recombinação entre três *loci* no milho são apresentadas aqui.

Loci	**Taxa de recombinação**
R e W_2	17%
R e L_2	35%
W_2 e L_2	18%

Qual é a ordem dos genes no cromossomo?

22. Nos tomates, o traço planta anã (*d*) é recessivo para planta alta (*D*), e folhas opacas (verde-claras) (*op*) é recessivo para folhas verdes (*Op*). Os *loci* que determinam a altura e a cor das folhas estão ligados e separados por uma distância de 7 m.u. Para cada um dos cruzamentos a seguir, determine os fenótipos e proporções dos descendentes produzidos.

a. $\frac{D}{d}\ \frac{Op}{op} \times \frac{d}{d}\ \frac{op}{op}$

b. $\frac{D}{d}\ \frac{op}{Op} \times \frac{d}{d}\ \frac{op}{op}$

c. $\frac{D}{d}\ \frac{Op}{op} \times \frac{D}{d}\ \frac{Op}{op}$

d. $\frac{D}{d}\ \frac{op}{Op} \times \frac{D}{d}\ \frac{op}{Op}$

23. Nas baratas-alemães (*Blatella germanica*), os olhos salientes (*bu*) são recessivos para os olhos normais (bu^+) e as asas curvas (*cv*) são recessivas para as asas retas (cv^+). Ambos os traços são codificados por genes autossômicos ligados. Uma barata-alemã tem o genótipo $bu^+\ bu\ cv^+\ cv$, mas os genes estão em repulsão. Quais dos conjuntos de genes a seguir serão encontrados nos gametas mais comuns produzidos por esta barata?
    a. $bu^+\ cv^+$
    b. *bu cv*
    c. $bu^+\ bu$
    d. $cv^+\ cv$
    e. $bu\ cv^+$

    Explique sua resposta.
24. Na *Drosophila melanogaster*, o corpo cor de ébano (*e*) e os olhos rugosos (*ro*) são codificados pelos genes autossômicos recessivos encontrados no cromossomo 3; eles estão separados por 20 m.u. O gene que codifica as cerdas bifurcadas (*f*) é recessivo ligado ao X e se separa independentemente de *e* e *ro*. Indique os fenótipos dos descendentes e suas proporções esperadas quando é feito um cruzamento-teste entre uma fêmea de cada um dos seguintes genótipos com um macho.

a. $\frac{e^+}{e}\ \frac{ro^+}{ro}\ \ \frac{f^+}{f}$

b. $\frac{e^+}{e}\ \frac{ro}{ro^+}\ \ \frac{f^+}{f}$

ANÁLISE DE DADOS

25. As abelhas têm determinação sexual haplodiploide: as fêmeas são diploides, se desenvolvendo a partir de óvulos fertilizados, enquanto os machos são haploides, se desenvolvendo a partir de óvulos não fertilizados. Otto Mackensen estudou as relações de ligação entre oito mutações nas abelhas (O. Mackensen. 1958. *Journal of Heredity* 49:99-102). O quadro a seguir indica os resultados dos dois cruzamentos de MacKensen, incluindo as três mutações recessivas: *cd* (corpo cor de cordovão), *h* (sem pelo) e *ch* (olhos cor verde-amarelada).

Genótipo da rainha	Fenótipos dos descendentes do zangão (macho)
$\frac{cd}{cd^+}\ \frac{h^+}{h}$	294 cordovões, 236 sem pelo, 262 cordovões e sem pelo, 289 tipo selvagem
$\frac{h}{h^+}\ \frac{ch^+}{ch}$	3.131 sem pelo, 3.064 verde-amareladas, 96 verde-amareladas e sem pelo, 132 tipo selvagem

    a. Apenas o genótipo da rainha é fornecido. Por que o genótipo do genitor não é necessário para mapear esses genes? Seria necessário o genótipo do genitor macho se examinarmos as descendentes fêmeas em vez dos descendentes machos?
    b. Determine os descendentes não recombinantes e recombinantes para cada cruzamento e calcule as distâncias de mapa entre

cd, *h* e *ch*. Desenhe um mapa de ligação ilustrando as disposições de ligação entre estes três genes.

26. Faça um teste qui-quadrado de independência nos dados fornecidos na Figura 7.2 para determinar se os genes para a cor da flor e formato do pólen nas ervilhas-de-cheiro se separam de forma independente. Indique o valor do qui-quadrado, graus de liberdade e probabilidade associada. Qual conclusão você tiraria sobre a segregação independente desses genes?

*27. Vários cruzamentos de dois pontos foram feitos entre sete *loci* (*a, b, c, d, e, f* e *g*), produzindo as frequências de recombinação a seguir. Faça o mapa dos sete *loci*, mostrando seus grupos de ligação, a ordem dos *loci* em cada grupo de ligação e as distâncias entre os *loci* de cada grupo.

Loci	Recombinação em porcentagem	*Loci*	Recombinação em porcentagem
a e *b*	50	*c* e *d*	50
a e *c*	50	*c* e *e*	26
a e *d*	12	*c* e *f*	50
a e *e*	50	*c* e *g*	50
a e *f*	50	*d* e *e*	50
a e *g*	4	*d* e *f*	50
b e *c*	10	*d* e *g*	8
b e *d*	50	*e* e *f*	50
b e *e*	18	*e* e *g*	50
b e *f*	50	*f* e *g*	50
b e *g*	50		

ANÁLISE DE DADOS

28. R. W. Allard e W. M. Clement determinaram as taxas de recombinação de alguns genes no feijão-de-lima (R. W. Allard and W. M. Clement. 1959. *Journal of Heredity* 50:63-67). O quadro a seguir lista as taxas de recombinação pareada para oito *loci* (*D, Wl, R, S, L_1, Ms, C,* e *G*) que eles mapearam. Com base nesses dados, desenhe alguns mapas genéticos para os diferentes grupos de ligação dos genes, indicando as distâncias entre os genes. Tenha em mente que essas taxas são estimativas das verdadeiras taxas de recombinação e que algum erro está associado a cada estimativa. Um asterisco ao lado de uma frequência de recombinação indica que ela é muito diferente de 50%.

Taxas de recombinação (%) entre sete loci no feijão-de-lima

	Wl	***R***	***S***	***L_1***	***Ms***	***C***	***G***
D	2,1*	39,3*	52,4	48,1	53,1	51,4	49,8
Wl		38,0*	47,3	47,7	48,8	50,3	50,4
R			51,9	52,7	54,6	49,3	52,6
S				26,9*	54,9	52,0	48,0
L_1					48,2	45,3	50,4
Ms						14,7*	43,1
C							52,0

*Significativamente diferente de 50%.

Seção 7.3

ANÁLISE DE DADOS

*29. Raymond Popp estudou a ligação entre os genes para olhos rosa (*p*), shaker-1 (*sh-1*, que provoca o comportamento rotatório, sacudidor de cabeça e surdez) e hemoglobina (*Hb*) em camundongos (R. A. Popp. 1962. *Journal of Heredity* 53:73-80). Ele fez vários cruzamentos-teste, nos quais camundongos heterozigotos para olhos rosa, shaker-1 e hemoglobina 1 e 2 foram cruzados com camundongos homozigotos para olhos rosa, shaker-1 e hemoglobina 2.

$$\frac{P\ Sh\text{-}1\ Hb^1}{p\ sh\text{-}1\ Hb^2} \times \frac{p\ sh\text{-}1\ Hb^2}{p\ sh\text{-}1\ Hb^2}$$

Foram produzidos os descendentes a seguir.

Genótipo dos descendentes	Número
$\frac{p\ sh\text{-}1\ Hb^2}{p\ sh\text{-}1\ Hb^2}$	274
$\frac{P\ Sh\text{-}1\ Hb^1}{p\ sh\text{-}1\ Hb^2}$	320
$\frac{P\ sh\text{-}1\ Hb^2}{p\ sh\text{-}1\ Hb^2}$	57
$\frac{p\ Sh\text{-}1\ Hb^1}{p\ sh\text{-}1\ Hb^2}$	45
$\frac{P\ Sh\text{-}1\ Hb^2}{p\ sh\text{-}1\ Hb^2}$	6
$\frac{p\ sh\text{-}1\ Hb^1}{p\ sh\text{-}1\ Hb^2}$	5
$\frac{p\ sh\text{-}1\ Hb^2}{p\ sh\text{-}1\ Hb^2}$	0
$\frac{p\ sh\text{-}1\ Hb^1}{p\ sh\text{-}1\ Hb^2}$	1
Total	708

a. Determine a ordem dos genes no cromossomo.
b. Calcule as distâncias do mapa entre os genes.
c. Determine o coeficiente de coincidência e a interferência para esses genes.

30. Endosperma macio (*wx*), endosperma encolhido (*sh*) e muda amarela (*v*) são codificados por três genes recessivos no milho, que estão ligados no cromossomo 5. Um milho homozigoto para todos os três alelos recessivos é cruzado com um milho homozigoto para todos os alelos dominantes. Os F_1 resultantes são cruzados com uma planta homozigota para os alelos recessivos em um cruzamento-teste de três pontos. Os descendentes do cruzamento-teste são:

wx	*sh*	*V*	87
Wx	*Sh*	*v*	94
Wx	*Sh*	*V*	3.479
wx	*sh*	*v*	3.478
Wx	*sh*	*V*	1.515
wx	*Sh*	*v*	1.531
wx	*Sh*	*V*	292
Wx	*sh*	*v*	280
Total			10.756

a. Determine a ordem dos genes no cromossomo.
b. Calcule as distâncias do mapa entre os genes.
c. Determine o coeficiente de coincidência e a interferência para esses genes.

31. Priscilla Lane e Margaret Green estudaram as relações das ligações de três genes que afetam a cor do pelo nos camundongos: mogno (*mg*), agouti (*a*) e desarrumado (*Rg*). Elas fizeram vários cruzamentos de três pontos, cruzando camundongos heterozigotos em todos os três *loci* com camundongos homozigotos para os alelos recessivos nesses *loci* (P W. Lane and M. C. Green. 1960. *Journal of Heredity* 51:228-230). O quadro a seguir lista os resultados dos cruzamentos-teste.

	Fenótipo		Número
a	*Rg*	+	1
+	+	*mg*	1
a	+	+	15
+	*Rg*	*mg*	9
+	+	+	16
a	*Rg*	*mg*	36
a	+	*mg*	76
+	*Rg*	+	69
Total			213

Nota: + representa um alelo do tipo selvagem.

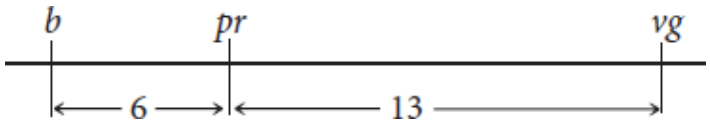
a. Determine a ordem dos *loci* que codificam os traços mogno, agouti e desarrumado no cromossomo, as distâncias no mapa entre eles e a interferência e coeficiente de coincidência para os genes.
b. Desenhe uma imagem dos dois cromossomos nos camundongos heterozigotos triplos usados nos cruzamentos-teste, indicando quais dos alelos estão presentes em cada cromossomo.

32. Espinhos finos (*s*), fruta lisa (*tu*) e cor uniforme da fruta (*u*) são três traços recessivos no pepino, genes esses ligados no mesmo cromossomo. Um pepino heterozigoto para os três traços é usado em um cruzamento-teste e os seguintes descendentes são produzidos a partir dele:

S	*U*	*Tu*	2
s	*u*	*Tu*	70
S	*u*	*Tu*	21
s	*u*	*tu*	4
S	*U*	*tu*	82
s	*U*	*tu*	21
s	*U*	*Tu*	13
S	*u*	*tu*	17
Total			230

a. Determine a ordem dos genes no cromossomo.
b. Calcule as distâncias do mapa entre os genes.
c. Determine o coeficiente de coincidência e a interferência para esses genes.
d. Liste os genes encontrados em cada cromossomo dos genitores usados no cruzamento-teste.

33. Na *Drosophila melanogaster*, o corpo preto (*b*) é recessivo para o corpo cinza (b^+), os olhos púrpura (*pr*) são recessivos para os olhos vermelhos (pr^+) e as asas vestigiais (*vg*) são recessivas para as asas normais (vg^+). Os *loci* que codificam esses traços estão ligados, com as seguintes distâncias do mapa:



A interferência entre esses genes é de 0,5. Uma mosca com corpo preto, olhos púrpura e asas vestigiais é cruzada com uma mosca homozigota para o corpo cinza, olhos vermelhos e asas normais. As fêmeas descendentes são cruzadas com machos com corpo preto, olhos púrpura e asas vestigiais. Se forem produzidos 1.000 descendentes desse cruzamento-teste, qual seriam os fenótipos e proporções encontrados?

34. *Olhos sépia*, *ausência de cerdas* e *tórax listrado* são três mutações recessivas na *Drosophila* encontradas no cromossomo 3. Um estudante de genética cruzou uma mosca homozigota para olhos sépia, ausência de cerdas e tórax listrado com uma mosca homozigota para os traços do tipo selvagem – olhos vermelhos, cerdas normais e tórax de cor uniforme. As fêmeas descendentes são, então, submetidas ao cruzamento-teste com machos com olhos sépia, ausência de cerdas e tórax listrado. Considere que a interferência entre esses genes é de 0,2 e que são produzidas 400 moscas descendentes a partir desse cruzamento-teste. Com base nas distâncias de mapa fornecidas na Figura 7.14, preveja os genótipos e as proporções dos descendentes resultantes do cruzamento-teste.
35. Oito sequências de DNA de diferentes indivíduos são apresentadas a seguir.

Posição do nucleotídio

	1				5				10				15			
Sequência 1…	T	C	T	G	G	A	T	C	A	T	C	A	C	A	T…	
Sequência 2…	A	C	A	G	C	A	T	C	A	T	T	A	C	G	T…	
Sequência 3…	T	C	A	G	G	A	T	C	A	T	T	A	C	A	T…	
Sequência 4…	T	C	A	G	G	A	T	C	A	T	T	A	C	A	T…	
Sequência 5…	A	C	A	G	C	A	T	C	A	T	T	A	C	G	T…	
Sequência 6…	T	C	T	G	G	A	T	C	A	T	C	A	C	A	T…	
Sequência 7…	T	C	A	G	G	A	T	C	A	T	T	A	C	A	T…	
Sequência 8…	A	C	A	G	C	A	T	C	A	T	T	A	C	G	T…	

a. Indique as posições dos nucleotídios de todos os polimorfismos de único nucleotídio (SNPs, posições dos nucleotídios onde os indivíduos variam e nas quais a base está presente) nessas sequências.
b. Quantos haplótipos diferentes (conjuntos de variantes ligadas) são encontrados nessas oito sequências?
c. Indique o haplótipo de cada sequência ao listar as bases específicas em cada posição variável naquele haplótipo. (Dica: ver Figura 20.8).

*36. Um grupo de geneticistas está interessado em identificar os genes que atuam na suscetibilidade à asma. Eles estudaram a herança dos marcadores genéticos em várias famílias com duas ou mais crianças asmáticas e descobriram uma associação entre a ocorrência ou não de asma e um marcador genético no braço curto do cromossomo 20, calculando um *lod score* de 2 para essa associação. O que esse *lod score* indica sobre os genes que influenciam a asma?

Seção 7.4

37. Foi criado um painel de linhagens celulares a partir de fusões de células somáticas de seres humanos e camundongos. Foram pesquisados cromossomos de humanos e a produção de uma enzima em cada linhagem. Foram obtidos os seguintes resultados:

Linhagem celular	Enzima	Cromossomos humanos											
		1	2	3	4	5	6	7	8	9	10	17	22
A	–	+	–	–	–	+	–	–	–	–	–	+	–
B	+	+	+	–	–	–	–	–	+	–	–	+	+
C	–	+	–	–	–	+	–	–	–	–	–	–	+
D	–	–	–	–	+	–	–	–	–	–	–	–	–
E	+	+	–	–	–	–	–	–	+	–	+	+	–

Com base nesses resultados, qual cromossomo tem o gene codificador da enzima?

*38. Foi criado um painel de linhagens celulares a partir de fusões de células somáticas de seres humanos e camundongos. Foram pesquisados cromossomos de humanos e a produção de três enzimas em cada linhagem. Foram obtidos os seguintes resultados:

Linhagem celular	Enzima	Cromossomos humanos										
	1	2	3	4	8	9	12	15	16	17	22	X
A	+	–	+	–	–	+	–	+	+	–	–	+
B	+	–	–	–	–	+	–	–	+	+	–	–
C	–	+	+	+	–	–	–	–	–	+	–	+
D	–	+	+	+	+	–	–	–	+	–	–	+

Com base nesses resultados, indique a localização cromossômica da enzima 1, enzima 2 e enzima 3.

*39. As localizações de seis deleções foram mapeadas no cromossomo da *Drosophila*, como indicado no mapa de deleção. Sabe-se que as mutações recessivas *a*, *b*, *c*, *d*, *e* e *f* estão localizadas na mesma região das deleções, mas a ordem das mutações no cromossomo não é conhecida.

Cromossomo

Deleção 1_____

Deleção 2_____

Deleção 3_____

Deleção 4_____

Deleção 5_____

Deleção 6_____

Quando as moscas homozigotas para as mutações recessivas são cruzadas com moscas homozigotas para as deleções, são obtidos os seguintes resultados, nos quais "m" representa um fenótipo mutante e um sinal de mais (+) representa o tipo selvagem. Com base nesses dados, determine a ordem relativa dos sete genes mutantes no cromossomo:

	Mutações					
Deleção	***a***	***b***	***c***	***d***	***e***	***f***
1	m	+	m	+	+	m
2	m	+	+	+	+	+
3	+	m	m	m	m	+
4	+	+	m	m	m	+
5	+	+	+	m	m	+
6	+	m	+	m	+	+

Questões desafiadoras

Seção 7.5

ANÁLISE DE DADOS

40. A transferrina é uma proteína do sangue codificada pelo *locus* da transferrina (*Trf*). Nos camundongos, os dois alelos nesse *locus* (Trf^a e Trf^b) são codominantes e codificam três tipos de transferrina:

Genótipo	**Fenótipo**
Trf^a/Trf^a	Trf-a
Trf^a/Trf^b	Trf-ab
Trf^b/Trf^b	Trf-b

O *locus* de diluição, encontrado no mesmo cromossomo, determina se a cor de um camundongo é diluída ou plena, sendo que um alelo para diluição (*d*) é recessivo para um alelo para cor uniforme (d^+):

Genótipo	**Fenótipo**
d^+d^+	d^+ (cor uniforme)
d^+d	d^+ (cor uniforme)
dd	d (diluição)

Donald Shreffler fez vários cruzamentos para determinar a distância de mapa entre o *locus* da transferrina e o *locus* de diluição (D. C. Shreffler. 1963. *Journal of Heredity* 54:127-129). O quadro a seguir apresenta vários cruzamentos feitos por Shreffler e os descendentes desses cruzamentos.

a. Calcule a frequência de recombinação entre os *loci Trf* e *d* ao usar os dados unificados de todos os cruzamentos.
b. Quais cruzamentos representam a recombinação na formação do gameta masculino e quais cruzamentos representam a recombinação na formação do gameta feminino?
c. Com base na sua resposta para a parte b, calcule a frequência de recombinação entre os genitores masculinos e femininos separadamente.
d. As taxas de recombinação nos machos e fêmeas são as mesmas? Se não forem, que fator seria responsável pela diferença?

Cruzamento	♂	♀	Fenótipos dos descendentes: d^+ Trf-ab	d^+ Trf-b	d Trf-ab	d Trf-b	Total
1	$\frac{d^+\ Trf^a}{d\ Trf^b}$	$\frac{d\ Trf^b}{d\ Trf^b}$	32	3	6	21	62
2	$\frac{d\ Trf^b}{d\ Trf^b}$	$\frac{d^+\ Trf^a}{d\ Trf^b}$	16	0	2	20	38
3	$\frac{d^+\ Trf^a}{d\ Trf^b}$	$\frac{d\ Trf^b}{d\ Trf^b}$	35	9	4	30	78
4	$\frac{d\ Trf^b}{d\ Trf^b}$	$\frac{d^+\ Trf^a}{d\ Trf^b}$	21	3	2	19	45
5	$\frac{d^+\ Trf^b}{d\ Trf^a}$	$\frac{d\ Trf^b}{d\ Trf^b}$	8	29	22	5	64
6	$\frac{d\ Trf^b}{d\ Trf^b}$	$\frac{d^+\ Trf^b}{d\ Trf^a}$	4	14	11	0	29

8 Variação Cromossômica

Como criar bananas de melhor qualidade

As bananas e as bananas-da-terra (coletivamente chamadas de bananas) são a fruta mais popular no mundo. Em países em desenvolvimento, elas são uma fonte de alimento importante, fornecendo amido e calorias a centenas de milhões de pessoas. Nos países industrializados, as bananas são as frutas mais consumidas; os americanos, por exemplo, consomem em quantidade equivalente às maçãs e laranjas combinadas. Mais de 100 milhões de toneladas de bananas são produzidas anualmente no mundo.

Não há uma distinção biológica concreta entre as bananas e as bananas-da-terra, mas o termo banana, em geral, refere-se às formas mais doces que são ingeridas cruas, enquanto banana-da-terra é usado para as que são descascadas não maduras e cozidas antes de comer. As bananas cultivadas são diferentes das espécies selvagens por não apresentarem sementes, o que as torna mais comestíveis mas impede sua reprodução. Os fazendeiros propagam as bananas de forma vegetativa, ao cortar partes das plantas existentes e induzi-las a crescer como novas plantas. Uma vez que elas se propagam de forma vegetativa, muitas bananas cultivadas são geneticamente idênticas.

Do ponto de vista genético, as bananas são interessantes porque muitas variedades têm conjuntos múltiplos de cromossomos. A maioria dos organismos eucariotos é diploide ($2n$), com dois conjuntos de cromossomos. Outros, como os fungos, são haploides (n), com um único conjunto de cromossomos. As bananas cultivadas são em muitos casos poliploides, com mais de dois conjuntos de cromossomos ($3n$, $4n$ ou mais). A maioria das cepas de bananas cultivadas foi criada ao cruzar as plantas dentro e entre duas espécies diploides: *Musa acuminata* (genoma = AA) e *Musa balbisiana* (genoma = BB). A maior parte das bananas cultivadas é triploide, com três conjuntos de cromossomos, tais como AAA, AAB ou ABB, e algumas bananas têm até quatro conjuntos de cromossomos (tetraploides), como AAAA, AAAB, AABB ou ABBB.

Apesar da sua importância como alimento, existem dificuldades para seu cultivo. A variante mais vendida nos hortifrutigranjeiros – a Cavendish – está ameaçada por doenças e pragas; nos últimos anos, um fungo de solo devastou plantações na Ásia. A antecessora da banana Cavendish, chamada de Gros Michel (Big Mike), era a preferida até uma doença dizimá-la entre os anos 1950 e 1960. Como a propagação vegetativa produz plantas geneticamente idênticas, as bananas cultivadas estão muito vulneráveis ao ataque de patógenos e pragas.

A fim de ajudar a desenvolver uma banana de melhor qualidade – mais resistente a doenças e pragas e mais nutritiva –, os geneticistas iniciaram um esforço internacional para sequenciar o genoma da fruta, o que produziu um esboço da sequência em 2012. Essa pesquisa demonstrou que o genoma da banana tem mais de 500 milhões de pares de bases (pb) de DNA e codifica 36.500 genes que produzem proteínas. Ao usar essa sequência de genoma, os cientistas já identificaram vários genes que atuam na resistência a doenças fúngicas, e estão explorando meios para criar e desenvolver bananas melhores por meio da engenharia genética.

A maioria das espécies apresenta um número característico de cromossomos, cada uma com tamanho e estrutura diferente, e todos os tecidos de um organismo (exceto os gametas) em geral têm o mesmo conjunto de cromossomos. Contudo, as variações no número de cromossomos – como os conjuntos extras de cromossomos observados nas bananas – surgem periodicamente. Também podem surgir variações na estrutura do cromossomo: os individuais podem perder ou ganhar partes, e a ordem dos genes em um cromossomo pode ser alterada. Essas variações no número e na estrutura dos cromossomos são chamadas **mutações cromossômicas** e exercem um papel importante na agricultura e na evolução da fruta.

Começaremos este capítulo com uma revisão rápida de alguns conceitos básicos da estrutura do cromossomo, que aprendemos no Capítulo 2. Consideraremos os diferentes tipos de mutações cromossômicas, suas definições, características, efeitos fenotípicos e influência na evolução.

8.1 Mutações cromossômicas incluem rearranjos aneuploides e

poliploides

Antes de considerarmos os diferentes tipos de mutações cromossômicas, seus efeitos e como eles surgem, revisaremos os fundamentos da estrutura do cromossomo.

Morfologia do cromossomo

Cada cromossomo funcional tem um centrômero ao qual as fibras do fuso se prendem e dois telômeros que estabilizam o cromossomo (ver Figura 2.7). Os cromossomos são classificados em quatro tipos básicos:

1. **Metacêntrico**. O centrômero está localizado aproximadamente no meio do cromossomo; logo, o cromossomo tem dois braços de mesmo tamanho.
2. **Submetacêntrico**. O centrômero está deslocado para uma extremidade, criando um braço longo e um curto. (Nos cromossomos humanos, o braço longo é indicado pela letra q, e o curto, pela letra p).
3. **Acrocêntrico**. O centrômero está próximo de uma extremidade, produzindo um braço longo, e um botão, ou satélite, na outra extremidade.
4. **Telocêntrico**. O centrômero está na extremidade do cromossomo ou muito próximo desta (ver Figura 2.8).

O conjunto completo de cromossomos que um organismo tem é chamado de *cariótipo*, e em geral é apresentado como uma imagem de cromossomos de metáfase alinhados na ordem decrescente do seu tamanho. Os cariótipos são preparados a partir de células em divisão ativa, como as células brancas do sangue, as da medula óssea ou as dos tecidos meristemáticos das plantas. Após o tratamento com uma substância química (como a colchicina) que impede que elas entrem na anáfase, as células são quimicamente preservadas, espalhadas em uma lâmina de microscópio, coradas e fotografadas. A fotografia é então ampliada, e os cromossomos individuais são isolados e organizados em um cariótipo. Para os cromossomos humanos, os cariótipos são preparados na rotina por máquinas automatizadas, que rastreiam uma lâmina usando uma câmera de vídeo presa a um microscópio, à procura de cromossomos dispersos. Quando uma dispersão é localizada, a câmera tira uma foto dos cromossomos, a imagem é digitalizada e os cromossomos são classificados e organizados eletronicamente por um computador.

As técnicas de preparo e coloração ajudam a distinguir os cromossomos de tamanho e formato semelhantes. Por exemplo, o preparo e a coloração especial de cromossomos com um corante especial chamado Giemsa revela as bandas G que distinguem as áreas de DNA ricas nos pares de bases adenina-timina (A-T) (**Figura 8.1 A**) (ver *Capítulo 10*). As bandas Q (**Figura 8.1 B**) são reveladas ao corar os cromossomos com mostarda de quinacrina e visualizar os cromossomos sob luz ultravioleta; a variação na luminosidade das bandas Q é o resultado das diferenças nas quantidades relativas de pares de bases de citosina-guanina (C-G) e adenina-timina (A-T). Outras técnicas relevam as bandas C (**Figura 8.1 C**), que são regiões do DNA ocupadas por heterocromatina centroméricas, e bandas R (**Figura 8.1 D**), ricas em pares de bases C-G.

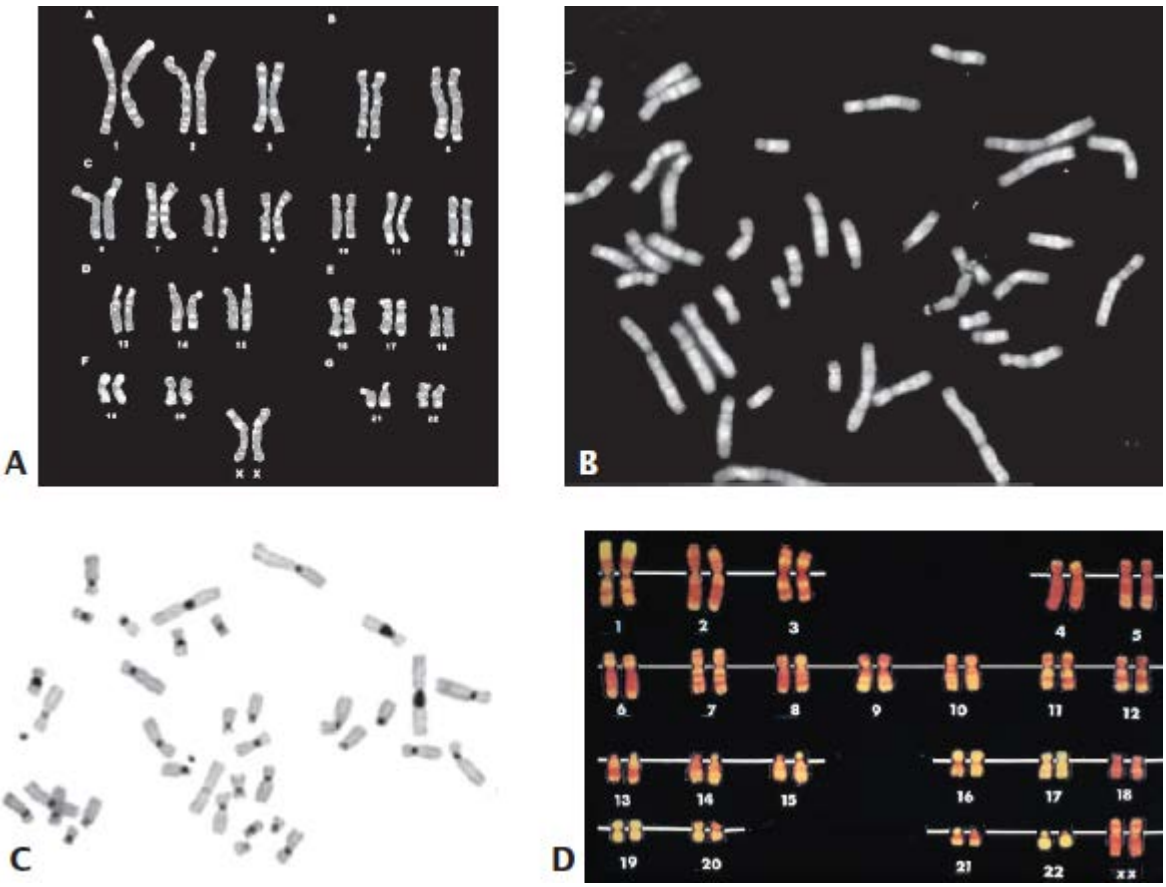


Figura 8.1 O bandeamento cromossômico é revelado por técnicas especiais de coloração. A. Bandeamento G. **B.** Bandeamento Q. **C.** Bandeamento C. **D.** Bandeamento R. (*Parte A*: Leonard Lessin/Science Source. *Partes B* e *C*: University of Washington Pathology Department. http://pathology.washington.edu. *Parte D*: Dr. Ram Verma/Phototake.)

Tipos de mutações cromossômicas

As mutações cromossômicas podem ser agrupadas em três categorias básicas: rearranjos cromossômicos, aneuploides e poliploides (**Figura 8.2**). Os rearranjos cromossômicos alteram a *estrutura* dos cromossomos; por exemplo, um segmento de cromossomo pode ser duplicado, eliminado ou invertido. Na aneuploidia, o *número* de cromossomos é alterado: um ou mais cromossomos são adicionados ou eliminados. Na poliploidia, um ou mais *conjuntos* de cromossomos são adicionados. Um poliploide é qualquer organismo com mais de dois conjuntos de cromossomos (3*n*, 4*n*, 5*n* ou mais).

8.2 Rearranjos cromossômicos modificam a estrutura dos cromossomos

Rearranjos cromossômicos são mutações que alteram as estruturas dos cromossomos. Os quatro tipos básicos de rearranjos são duplicações, deleções, inversões e translocações (**Figura 8.3**). Muitos deles se originam quando ocorrem rupturas das fitas duplas nas moléculas de DNA encontradas em um cromossomo. Uma vez que essas rupturas das fitas duplas no DNA em geral levam à morte, os organismos desenvolveram mecanismos elaborados para reparar tais rupturas ao reconectar as extremidades rompidas do DNA (ver *Capítulo 18*). Se as duas extremidades rompidas forem reunidas corretamente, o cromossomo original é restaurado e não há rearranjo cromossômico. Entretanto, às vezes as extremidades erradas são conectadas, levando ao rearranjo cromossômico. Os rearranjos cromossômicos também podem surgir por meio de erros no *crossing over* ou quando este ocorre entre sequências de DNA repetidas.

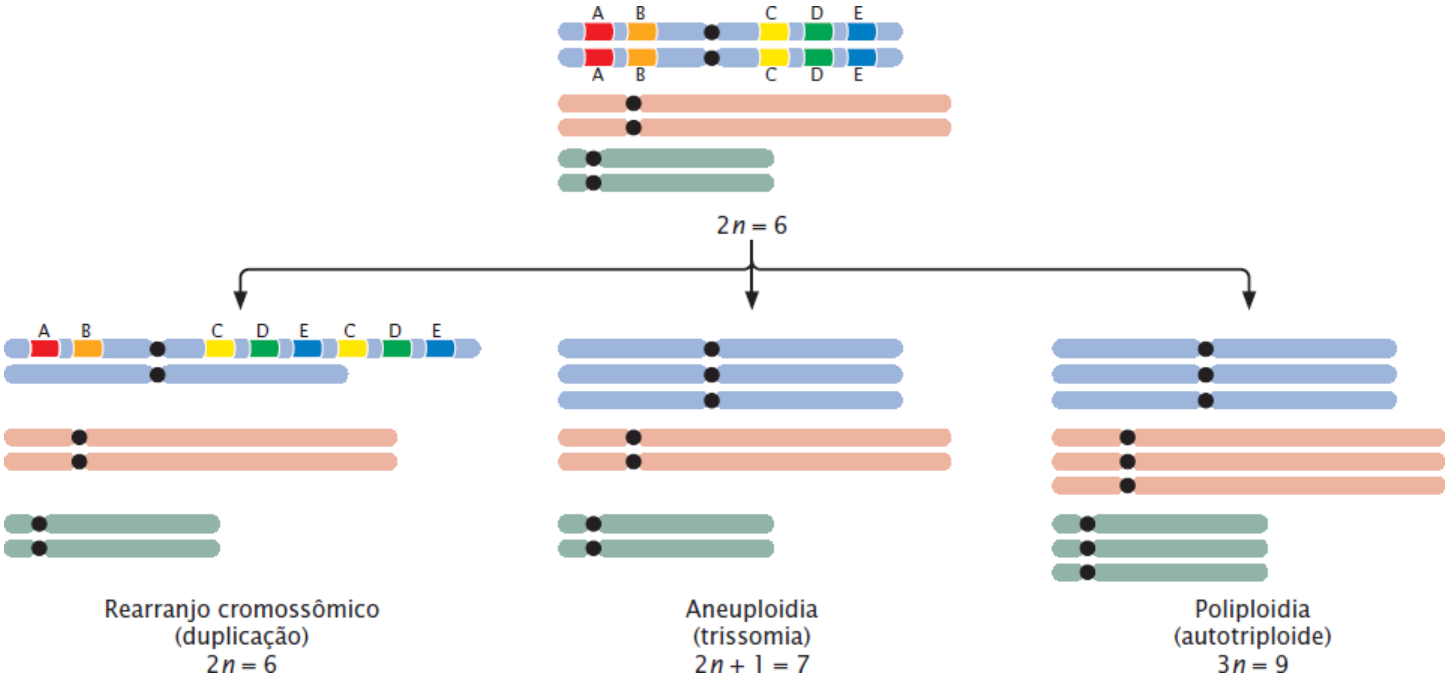


Figura 8.2 As mutações cromossômicas consistem em rearranjos cromossômicos, aneuploides e poliploides. Duplicações, trissomia e autotriploides são exemplos de cada tipo de mutação.

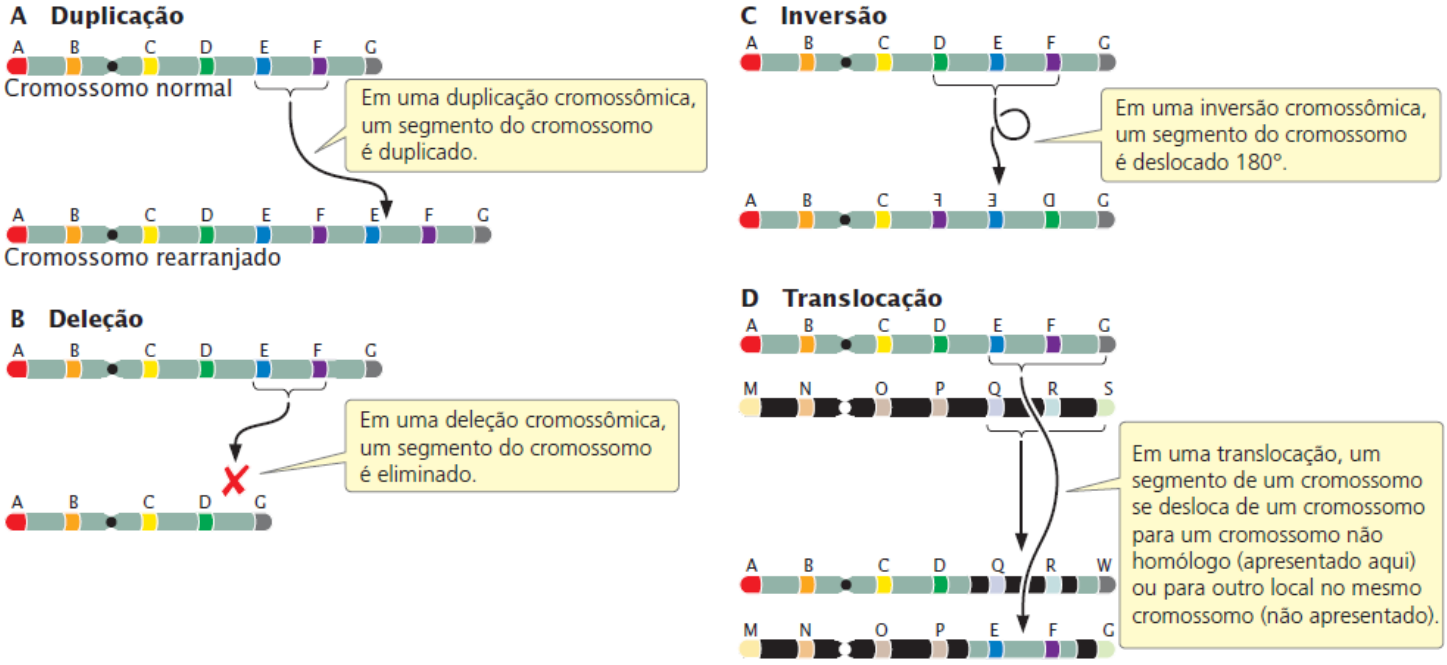


Figura 8.3 Os quatro tipos básicos de rearranjos cromossômicos são duplicações, deleções, inversões e translocações.

Duplicações

Uma **duplicação cromossômica** é uma mutação na qual parte do cromossomo é duplicada (ver Figura 8.3 A). Considere um cromossomo com o segmento AB•CDEFG, em que • representa o centrômero. Uma duplicação pode incluir os segmentos EF, dando origem a um cromossomo com os segmentos AB•CDEFEFG. Esse tipo de duplicação, na qual a região duplicada está imediatamente adjacente ao segmento original é chamada **duplicação em *tandem***. Se o segmento duplicado está localizado a uma distância do segmento original, no mesmo cromossomo ou em um cromossomo diferente, o rearranjo é chamado **duplicação deslocada**. Um exemplo de duplicação deslocada seria AB•CDEFGEF. Uma duplicação pode estar na mesma orientação da sequência original, como nos dois primeiros exemplos, ou invertida: AB•CDEFFEG. Nesse caso, ela é chamada **duplicação reversa**.

Efeitos das duplicações cromossômicas. Um indivíduo homozigoto para uma duplicação carreia a duplicação em ambos os cromossomos homólogos, e um indivíduo heterozigoto para uma duplicação tem um cromossomo normal e o outro com a duplicação. Nos heterozigotos (**Figura 8.4 A**), os problemas surgem no pareamento de cromossomos na prófase I da meiose, porque os dois cromossomos não são homólogos por toda sua extensão. O pareamento e a sinapse das regiões

homólogas requerem que um ou ambos os cromossomos estejam presos e girem de modo que essas regiões sejam capazes de se alinhar (**Figura 8.4 B**). A aparência dessa estrutura em alça característica na meiose é um meio de detectar duplicações.

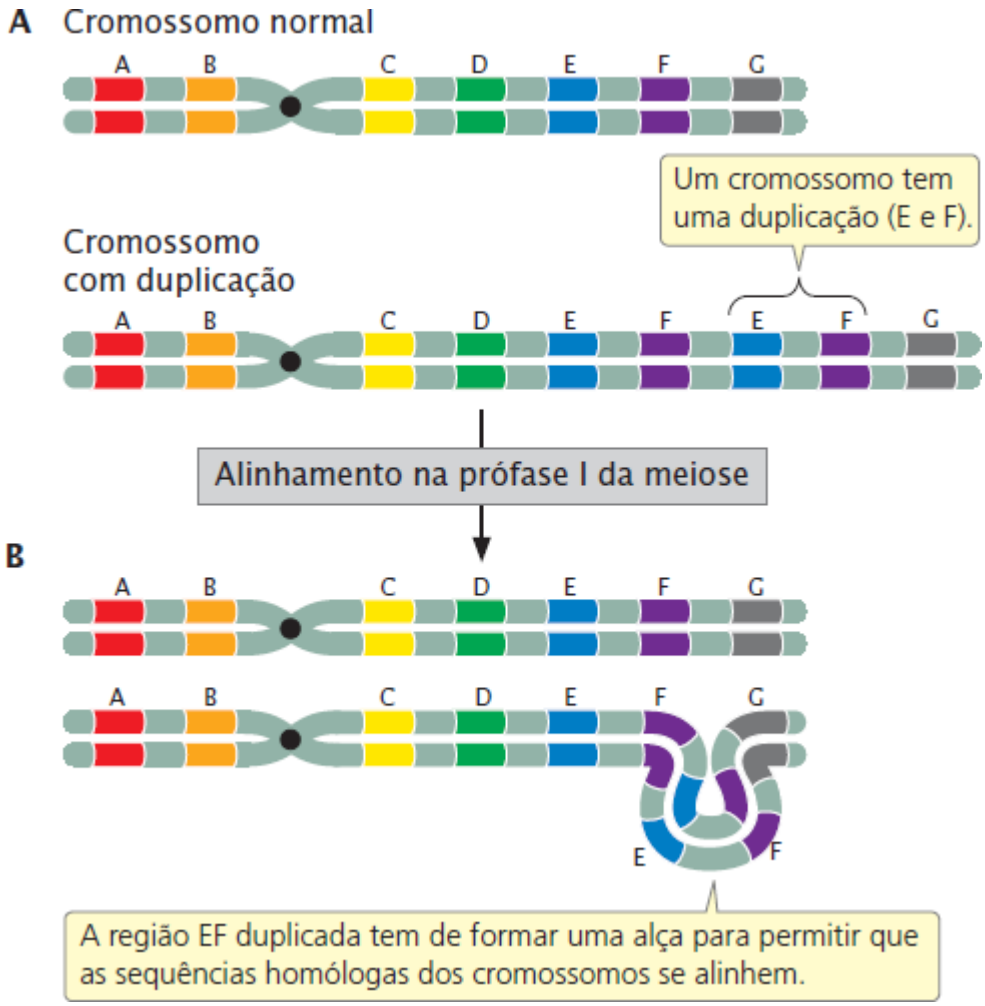


Figura 8.4 Em um indivíduo heterozigoto para uma duplicação, o cromossomo duplicado forma uma alça durante o pareamento na prófase I.

As duplicações podem ter efeitos importantes no fenótipo. Entre as moscas-da-fruta, por exemplo, uma com mutação *Bar* tem um número reduzido de facetas no olho, tornando-o menor e com formato de barra em vez do formato oval (**Figura 8.5**). A mutação *Bar* surge a partir de uma pequena duplicação no cromossomo X, herdado como um traço ligado ao X, dominante incompleto: as moscas-fêmeas heterozigotas têm olhos menores (o número de facetas é reduzido, ver Figura 8.5 B), enquanto nas moscas-fêmeas homozigotas e nos machos hemizigotos o número de facetas está bem reduzido (ver Figura 8.5 C). Em alguns casos, uma mosca carreia três cópias da duplicação *Bar* no seu cromossomo X; para as moscas que carreiam tais mutações, que são chamadas de *Bar dupla*, o número de facetas é bem reduzido (ver Figura 8.5 D).

As duplicações e deleções surgem a partir de um *crossing over* desigual, no qual os segmentos duplicados dos cromossomos não se alinham durante o processo. O *crossing* desigual é frequentemente a causa de daltonismo nos seres humanos. A percepção da cor é afetada pelos genes da opsina vermelha e verde, que são encontrados no cromossomo X e têm 98% da sua sequência de DNA idêntica. A maioria das pessoas com visão colorida normal tem um gene para opsina vermelha e outro para opsina verde (embora algumas pessoas tenham mais de uma cópia de cada). Às vezes, dois cromossomos X pareados em uma fêmea não se alinham de forma adequada na prófase I e ocorre um *crossing over* desigual. Esse *crossing over* desigual produz um cromossomo com um gene extra para opsina e outro sem esse gene (**Figura 8.6**; ver também Figura 18.12). Quando um homem herda o cromossomo sem os genes para opsina, ele é daltônico.

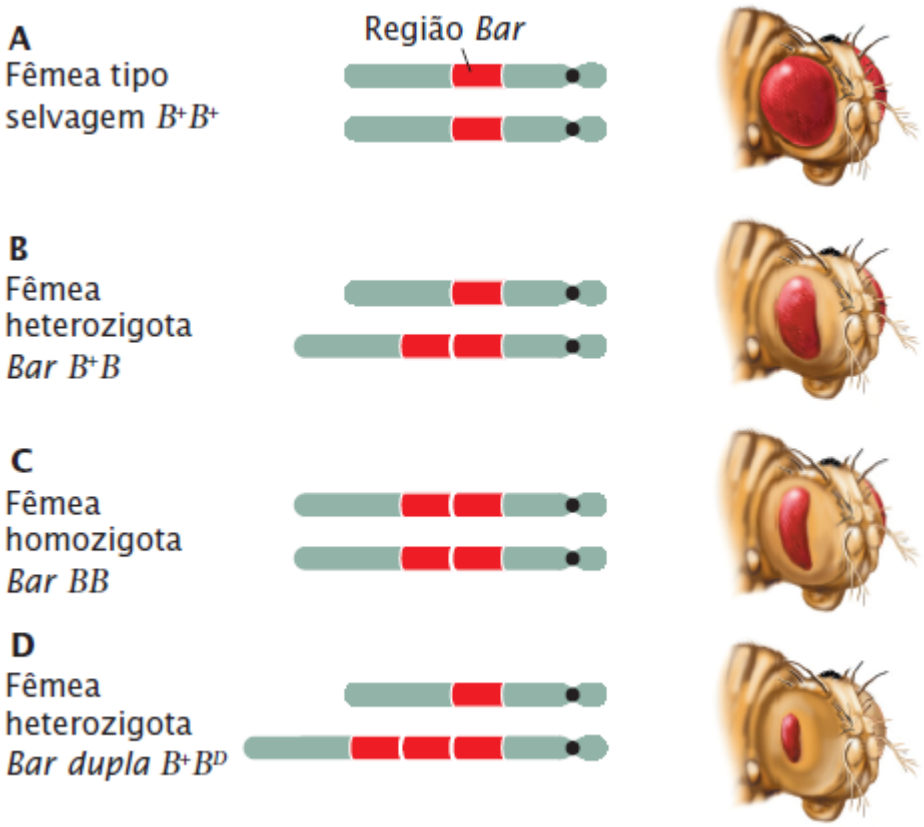


Figura 8.5 O fenótipo *Bar* na *Drosophila melanogaster* resulta de uma duplicação ligada ao X. (**A**) As moscas-da-fruta do tipo selvagem têm olhos de tamanho normal. (**B**) Moscas heterozigotas e (**C**) homozigotas para a mutação *Bar* têm olhos menores, em formato de barra. (**D**) As moscas com *Bar dupla* têm três cópias da duplicação e olhos muito menores em formato de barra.

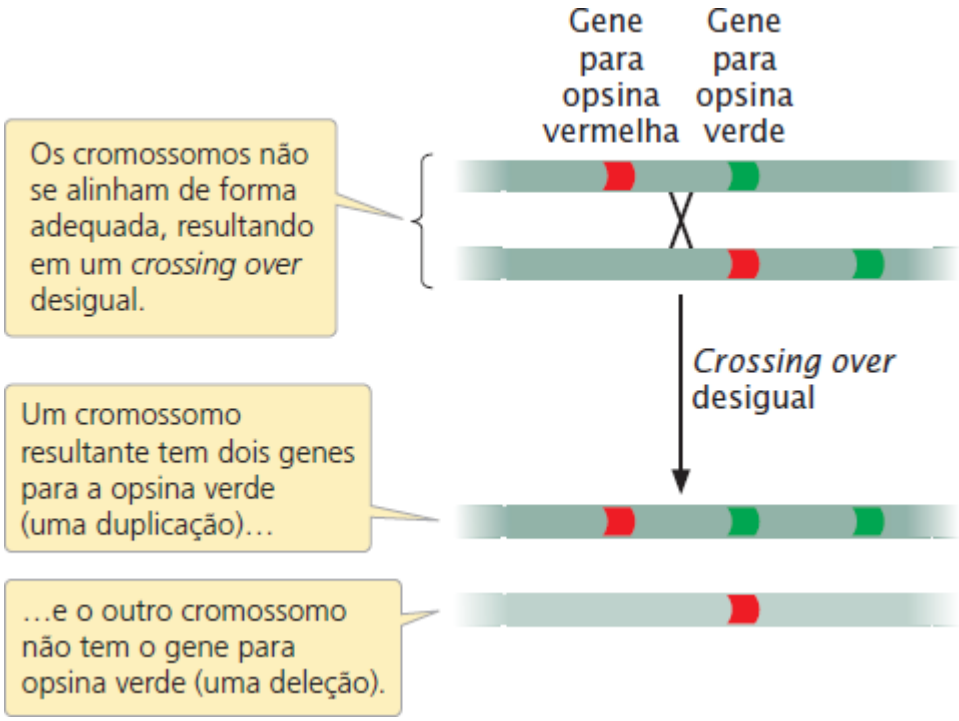


Figura 8.6 O *crossing over* desigual produz duplicações e deleções.

Dosagem gênica não balanceada. Como uma duplicação cromossômica altera o fenótipo? Apesar de tudo, as sequências de genes não são alteradas pelas duplicações e nenhuma informação genética é perdida; a única alteração é a existência de cópias adicionais de sequências normais. A resposta para essa pergunta não está bem compreendida, mas os efeitos são causados mais provavelmente pelo desequilíbrio na quantidade dos produtos gênicos (dosagem gênica anormal). A quantidade de uma proteína sintetizada por uma célula está diretamente relacionada com o número de cópias dos seus genes correspondentes: em geral, um organismo com três cópias funcionais de um gene produz 1,5 vez mais da proteína codificada por aquele gene do que a proteína produzida por um indivíduo com duas cópias. Como os processos de desenvolvimento exigem a interação de muitas proteínas, em geral elas dependem de forma crítica da dosagem gênica adequada. Se a quantidade

de uma proteína aumentar enquanto a das outras permanece constante, podem surgir problemas (**Figura 8.7**). As duplicações podem ter graves consequências quando o equilíbrio preciso de um produto gênico é crítico para o funcionamento da célula (**Quadro 8.1**).

Duplicações segmentares. O genoma humano tem muitas sequências duplicadas chamadas **duplicações segmentares**, que são definidas como duplicações maiores que 1.000 pares de bases (1.000 pb) em comprimento. Na maioria das duplicações segmentares, as duas cópias são encontradas no mesmo cromossomo (uma duplicação intracromossômica), mas, em outras, as duas cópias são encontradas em diferentes cromossomos (uma duplicação intercromossômica). Muitas duplicações segmentares têm sido detectadas com o uso de técnicas moleculares que examinam as sequências de DNA em um cromossomo (ver *Capítulo 20*). Essas técnicas revelam que cerca de 4% do genoma humano têm duplicações segmentares. No genoma humano, o tamanho médio das duplicações segmentares é de 15.000 pb.

Importância das duplicações na evolução. As duplicações têm surgido frequentemente ao longo da evolução de muitos organismos eucariotos. As duplicações cromossômicas fornecem uma forma na qual novos genes evoluem. Em muitos casos, as cópias existentes de um gene não estão livres para variar, porque elas codificam um produto que é essencial para o desenvolvimento ou função. Entretanto, após um cromossomo sofrer duplicação, existem cópias extras do gene na região duplicada. A cópia original pode fornecer a função essencial, enquanto uma cópia extra a partir da duplicação está livre para sofrer mutações e alterações. Com a evolução, a cópia extra pode adquirir mutações suficientes para assumir uma nova função que beneficie o organismo. Por exemplo, os seres humanos têm vários genes que codificam diferentes cadeias de globina, algumas das quais funcionam como carreadoras de oxigênio durante os estágios adultos, e outras que funcionam durante o desenvolvimento embrionário e fetal. Todos esses genes para globinas surgiram a partir de um gene ancestral original que sofreu várias duplicações.

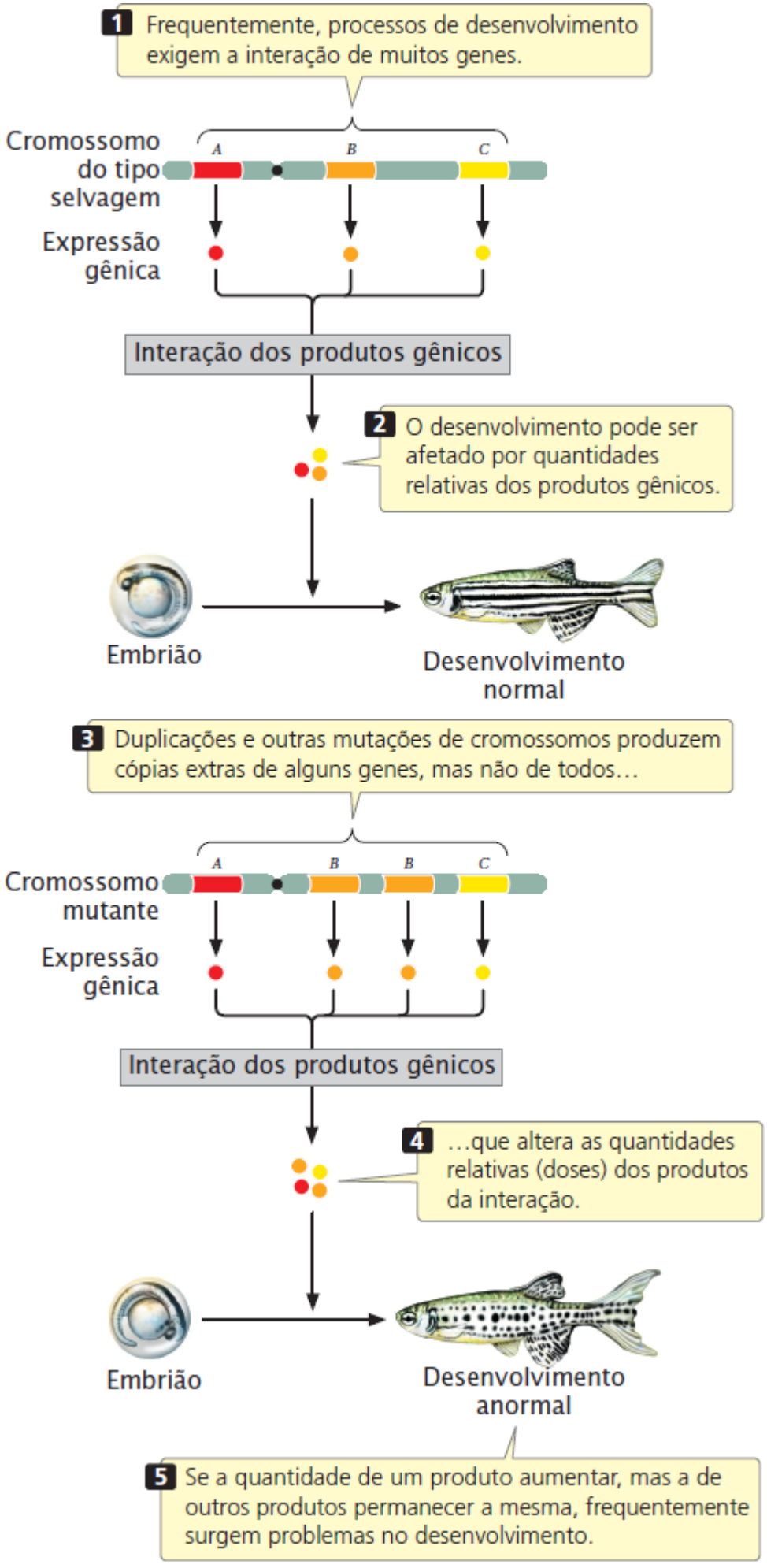


Figura 8.7 A dosagem gênica não balanceada leva a anomalias no desenvolvimento.

Quadro 8.1 Efeitos de alguns rearranjos dos cromossomos humanos.

Tipo de rearranjo	Cromossomo	Distúrbios	Sintomas
Duplicação	4, braço curto	–	Cabeça pequena, pescoço curto, linha de implantação dos cabelos baixa, retardo do crescimento e intelectual
Duplicação	4, braço longo	–	Cabeça pequena, testa inclinada, anomalias nas mãos
Duplicação	7, braço longo	–	Atraso no desenvolvimento, assimetria da cabeça, couro cabeludo com cabelo ralo, nariz pequeno, orelhas de implantação baixa
Duplicação	9, braço curto	–	Rosto característico, déficit intelectual variável, testa alta e ampla, anomalias nas mãos
Deleção	5, braço curto	Síndrome do miado do gato	Cabeça pequena, choro diferente, olhos muito afastados, rosto redondo, déficit intelectual
Deleção	4, braço curto	Síndrome de Wolf-Hirschhorn	Cabeça pequena com testa elevada, nariz largo, fenda labiopalatina, grave déficit intelectual
Deleção	4, braço longo	–	Cabeça pequena, déficit intelectual leve a moderado, fenda labiopalatina, anomalias nas mãos e nos pés
Deleção	7, braço longo	Síndrome de Williams-Beuren	Características faciais, defeitos cardíacos, déficit mental
Deleção	15, braço longo	Síndrome de Prader-Willi	Dificuldade para se alimentar no início, mas se torna obeso após 1 ano de vida, déficit intelectual leve a moderado
Deleção	18, braço curto	–	Rosto redondo, grandes orelhas de implantação baixa, déficit intelectual leve a moderado
Deleção	18, braço longo	–	Formato diferente da boca, mãos pequenas, cabeça pequena, déficit intelectual

Conceitos

Uma duplicação cromossômica é uma mutação que duplica parte de um cromossomo. Em indivíduos heterozigotos para uma duplicação cromossômica, a região duplicada do cromossomo forma uma alça quando os cromossomos homólogos se pareiam na prófase I da meiose. As duplicações podem ter efeitos significativos no fenótipo, possivelmente ao alterar a dosagem gênica. As duplicações segmentares são comuns no genoma humano e têm um papel importante na evolução de muitos eucariotos.

Checagem dos conceitos 1

As duplicações segmentares resultam em fenótipos anormais porque:

a. Os processos de desenvolvimento dependem de quantidades relativas de proteínas codificadas por diferentes genes.
b. Cópias extras de genes na região duplicada não pareiam na meiose.
c. É mais provável que o cromossomo se rompa quando ele forma uma alça na meiose.
d. O DNA extra tem de ser replicado, o que reduz a velocidade da divisão celular.

Deleções

Um segundo tipo de rearranjo cromossômico é a **deleção cromossômica**: a perda de um segmento de cromossomo (ver Figura 8.3 B). Um cromossomo com os segmentos AB•CDEFG que sofre uma deleção do segmento EF pode gerar um cromossomo AB•CDG mutante.

Uma grande deleção pode ser facilmente detectada porque o cromossomo fica flagrantemente mais curto. Em indivíduos heterozigotos para deleções, o cromossomo normal precisa formar uma alça durante o pareamento dos homólogos na prófase I da meiose (**Figura 8.8**) para permitir que as regiões homólogas dos dois cromossomos se alinhem e sofram sinapse. Essa alça gera uma estrutura muito parecida com a observada para indivíduos heterozigotos para duplicações.

Efeitos das deleções. As consequências fenotípicas de uma deleção dependem de quais genes estão localizados na região eliminada. Se a deleção incluir o centrômero, então o cromossomo não se separa na meiose ou na mitose e será perdido. Muitas deleções são letais no homozigoto porque todas as cópias de qualquer gene essencial localizado na região eliminada são perdidas. Até indivíduos heterozigotos para uma deleção podem ter múltiplos defeitos por três motivos.

Primeiro, a condição do heterozigoto pode produzir desequilíbrios na quantidade dos produtos do gene, semelhante aos desequilíbrios produzidos por cópias extras de gene. Segundo, normalmente, as mutações recessivas no cromossomo homólogo sem deleção podem ser expressas quando o alelo do tipo selvagem é eliminado (e não está mais presente para mascarar a expressão do alelo recessivo). A expressão de uma mutação recessiva é chamada de **pseudodominância**, e é uma indicação de que um dos cromossomos homólogos tem uma deleção. Terceiro, alguns genes têm de existir nas duas cópias para que haja função normal. Quando uma única cópia de um gene não é suficiente para produzir um fenótipo do tipo selvagem, ele é chamado **gene haploinsuficiente**. Uma série de mutações de asa ligadas ao X na *Drosophila* é conhecida como *Notch*. Essas mutações são, em muitos casos, resultantes de deleções cromossômicas. As deleções *Notch* se comportam como mutações dominantes: quando heterozigota para uma deleção *Notch*, uma mosca tem asas cortadas nas pontas e ao longo das bordas. A mutação *Notch* é, portanto, haploinsuficiente. As fêmeas homozigotas para uma deleção *Notch* (ou machos que são hemizigotos) morrem de forma prematura durante o desenvolvimento embrionário. O *locus Notch*, que é eliminado nas mutações *Notch*, codifica um receptor que normalmente transmite sinais recebidos de fora da célula para seu interior, e é importante para o desenvolvimento da mosca. A deleção atua como um gene recessivo letal porque a perda de todas as cópias do gene *Notch* impede o desenvolvimento normal.

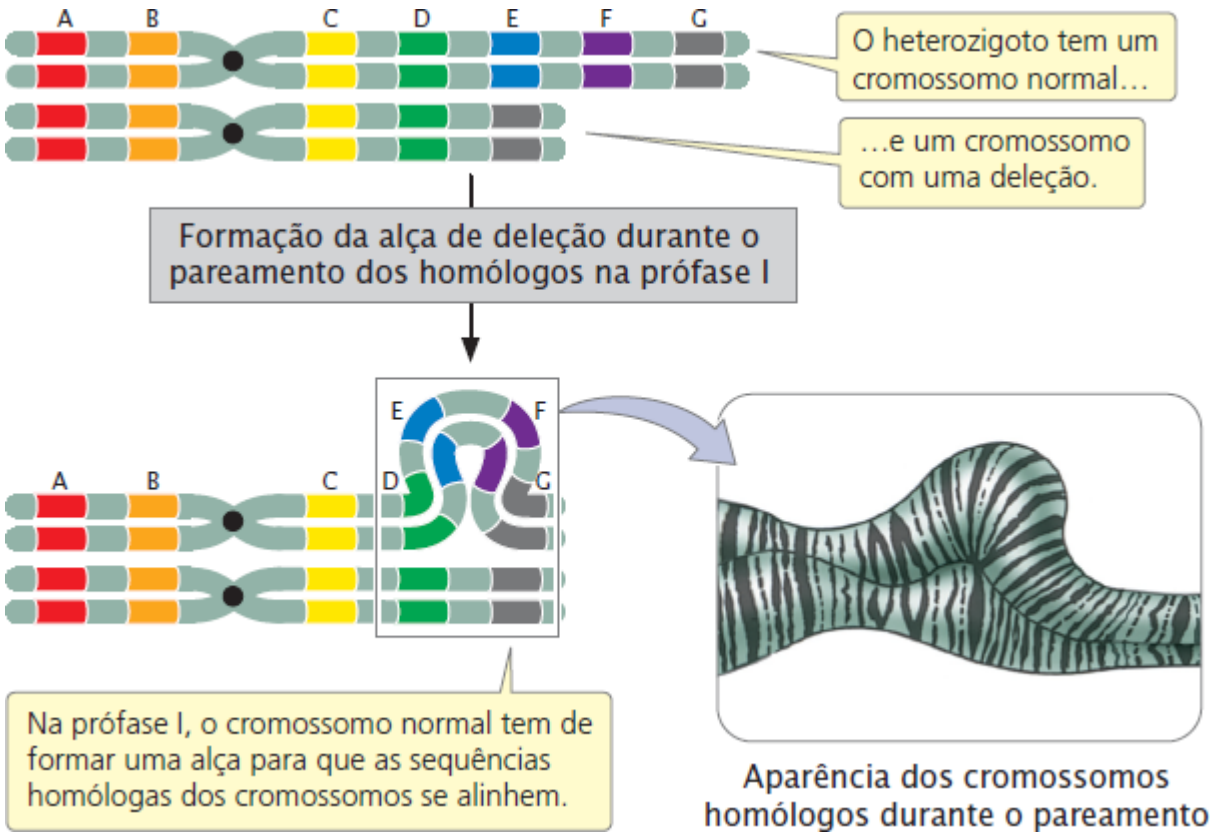


Figura 8.8 Em um indivíduo heterozigoto para uma deleção, o cromossomo normal forma uma alça durante o pareamento na prófase I.

Deleções cromossômicas nos seres humanos. Nos seres humanos, uma deleção no braço curto do cromossomo 5 é responsável pela síndrome do miado do gato ou *cri-du-chat*. O nome (do francês, "miado de gato") é originado pelo som emitido, semelhante ao miado de um gato, pelos recém-nascidos com essa síndrome. Uma criança heterozigota para essa deleção tem cabeça pequena, olhos bem espaçados, rosto redondo e déficit intelectual. A deleção de parte do braço curto do cromossomo 4 resulta em outro distúrbio humano, a síndrome de Wolf-Hirschhorn, caracterizado por convulsões, grave déficit intelectual e atraso no crescimento. Uma deleção de um minúsculo segmento do cromossomo 7 causa haploinsuficiência do gene que codifica a elastina e alguns outros genes e leva a uma condição conhecida como síndrome de Williams-Beuren, caracterizada por traços faciais distintos, defeitos cardíacos, hipertensão arterial e comprometimento cognitivo. Os efeitos das deleções nos seres humanos estão resumidos no Quadro 8.1.

Conceitos

Uma deleção cromossômica é uma mutação na qual uma parte de um cromossomo é perdida. Nos indivíduos heterozigotos para uma deleção, o cromossomo normal forma uma alça durante o pareamento na prófase I da meiose. As deleções fazem com que os genes recessivos no cromossomo homólogo sejam expressos e podem causar desequilíbrios nos produtos do gene.

Checagem dos conceitos 2

O que é pseudodominância e como ela é produzida a partir de uma deleção cromossômica?

Inversões

Um terceiro tipo de rearranjo cromossômico é uma **inversão cromossômica**, na qual um segmento do cromossomo é invertido – gira 180° (ver Figura 8.3 C). Se um cromossomo originalmente tem os segmentos AB•CDEFG, então o cromossomo AB•CFEDG representa uma inversão que inclui os segmentos DEF. Para que ocorra uma inversão, o cromossomo deve ser quebrado em dois locais. As inversões que não incluem o centrômero, como AB•CFEDG, são chamadas de **inversões paracêntricas**, enquanto as inversões que incluem o centrômero, como ADC•BE FG, são chamadas **inversões pericêntricas**.

Os heterozigotos de inversões são comuns em muitos organismos, incluindo diversas plantas, algumas espécies de *Drosophila*, mosquitos e gafanhotos. As inversões podem ter um papel importante na evolução humana: os padrões de bandeamento G revelam que vários cromossomos humanos são diferentes dos chimpanzés por apenas uma inversão pericêntrica

(**Figura 8.9**).

Efeitos das inversões. Os organismos individuais com inversões não perdem nem ganham nenhum material genético; apenas a sequência de DNA foi alterada. Contudo, essas mutações têm profundos efeitos no fenótipo. Uma inversão pode romper o gene em duas partes, com uma delas se movendo para um novo local e destruindo a função daquele gene. Mesmo quando o cromossomo se rompe entre os genes, os efeitos fenotípicos podem surgir a partir da ordem do gene invertido em uma inversão. Muitos genes são regulados de forma dependente da posição; se suas posições são alteradas por uma inversão, sua expressão pode ser alterada, uma mudança chamada de **efeito da posição**. Por exemplo, quando uma inversão desloca um alelo do tipo selvagem (que normalmente codifica os olhos vermelhos) no *locus* branco na *Drosophila* para uma região do cromossomo que tem uma cromatina altamente condensada e inativa, o alelo do tipo selvagem não é expresso em algumas células, resultando em um olho com pontos vermelhos e brancos.

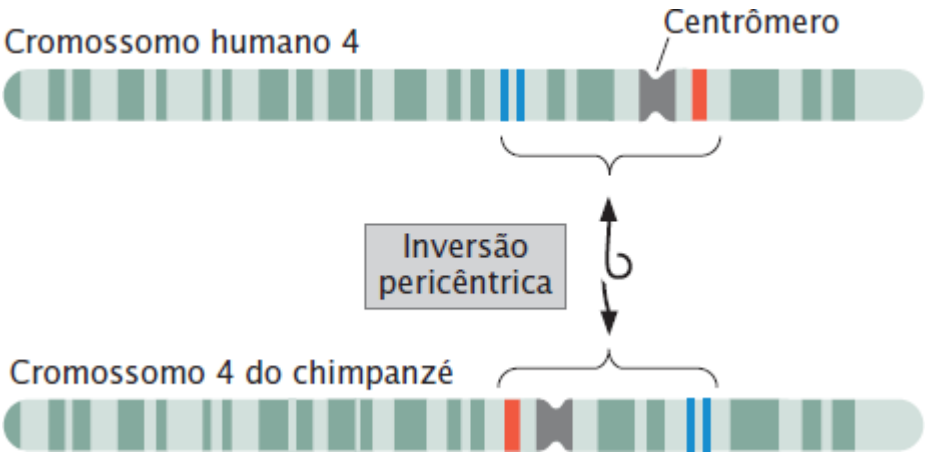


Figura 8.9 O cromossomo 4 é diferente entre os seres humanos e os chimpanzés por uma inversão pericêntrica. (Cortesia de Dra. Christine Harrison.)

Inversões na meiose. Quando um indivíduo é homozigoto para uma inversão não surgem problemas específicos na meiose, e os dois cromossomos homólogos podem parear e se separar normalmente. Entretanto, quando um indivíduo é heterozigoto para uma inversão, a ordem do gene dos dois homólogos é diferente e as sequências homólogas só conseguem se alinhar e parear se os dois cromossomos formarem uma alça de inversão (**Figura 8.10**).

Os indivíduos heterozigotos para inversões também exibem uma recombinação reduzida entre os genes localizados na região invertida. A frequência de *crossing over* na inversão não é reduzida de fato, mas, quando o *crossing over* ocorre, o resultado são gametas anormais que geram descendentes inviáveis e, nesse caso, não são observados descendentes recombinantes. Vamos ver como isso acontece.

A **Figura 8.11** ilustra os resultados do *crossing over* em uma inversão paracêntrica: o indivíduo é heterozigoto para uma inversão (ver Figura 8.11 A), com um cromossomo do tipo selvagem, não mutante (AB•CDEFG) e um cromossomo invertido (AB•EDCFG). Na prófase I da meiose, uma alça de inversão se forma, permitindo que as sequências homólogas se pareiem (ver Figura 8.11 B). Se um *crossing over* único ocorrer na região invertida (entre os segmentos C e D na Figura 8.11), surge uma estrutura incomum (ver Figura 8.11 C). As outras duas cromátides externas, que não participam no *crossing over*, apresentam sequências de genes originais, não recombinantes. As duas cromátides internas, que participam no *crossing over*, são muito anormais: cada uma tem duas cópias de alguns genes e nenhuma cópia de outros. Além disso, uma das quatro cromátides agora tem dois centrômeros e é chamada de **cromátide dicêntrica**; a outra não tem um centrômero e é uma **cromátide acêntrica**.

Na anáfase I da meiose, os centrômeros são puxados para os polos opostos e os dois cromossomos homólogos se separam. Essa ação alonga a cromátide dicêntrica ao longo do centro do núcleo, formando uma estrutura chamada **ponte dicêntrica** (ver Figura 8.11 D). Consequentemente, a ponte dicêntrica se rompe à medida que os dois centrômeros são puxados para sentidos opostos. As fibras do fuso não se prendem ao fragmento acêntrico, então este fragmento não se separa para um polo do fuso e em geral é perdido quando o núcleo se forma novamente.

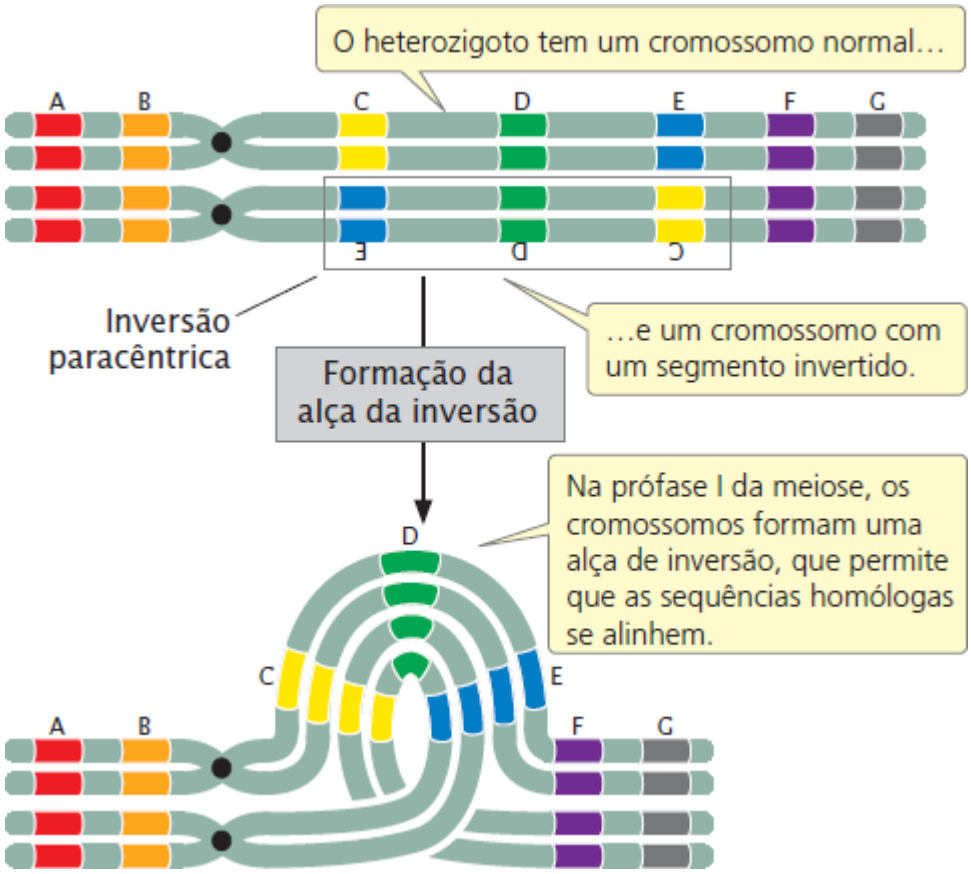
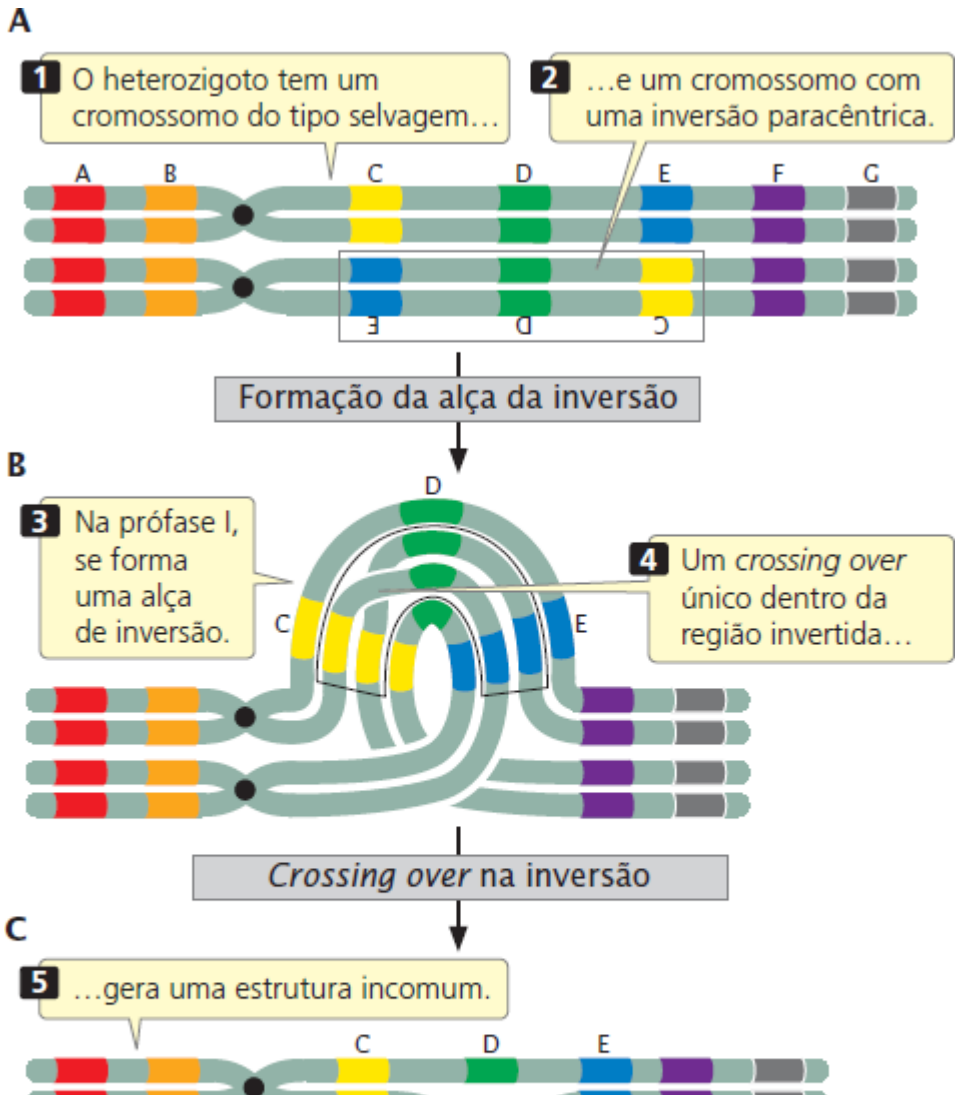


Figura 8.10 **Em um indivíduo heterozigoto para uma inversão paracêntrica, os cromossomos formam uma alça de inversão durante o pareamento na prófase I.**



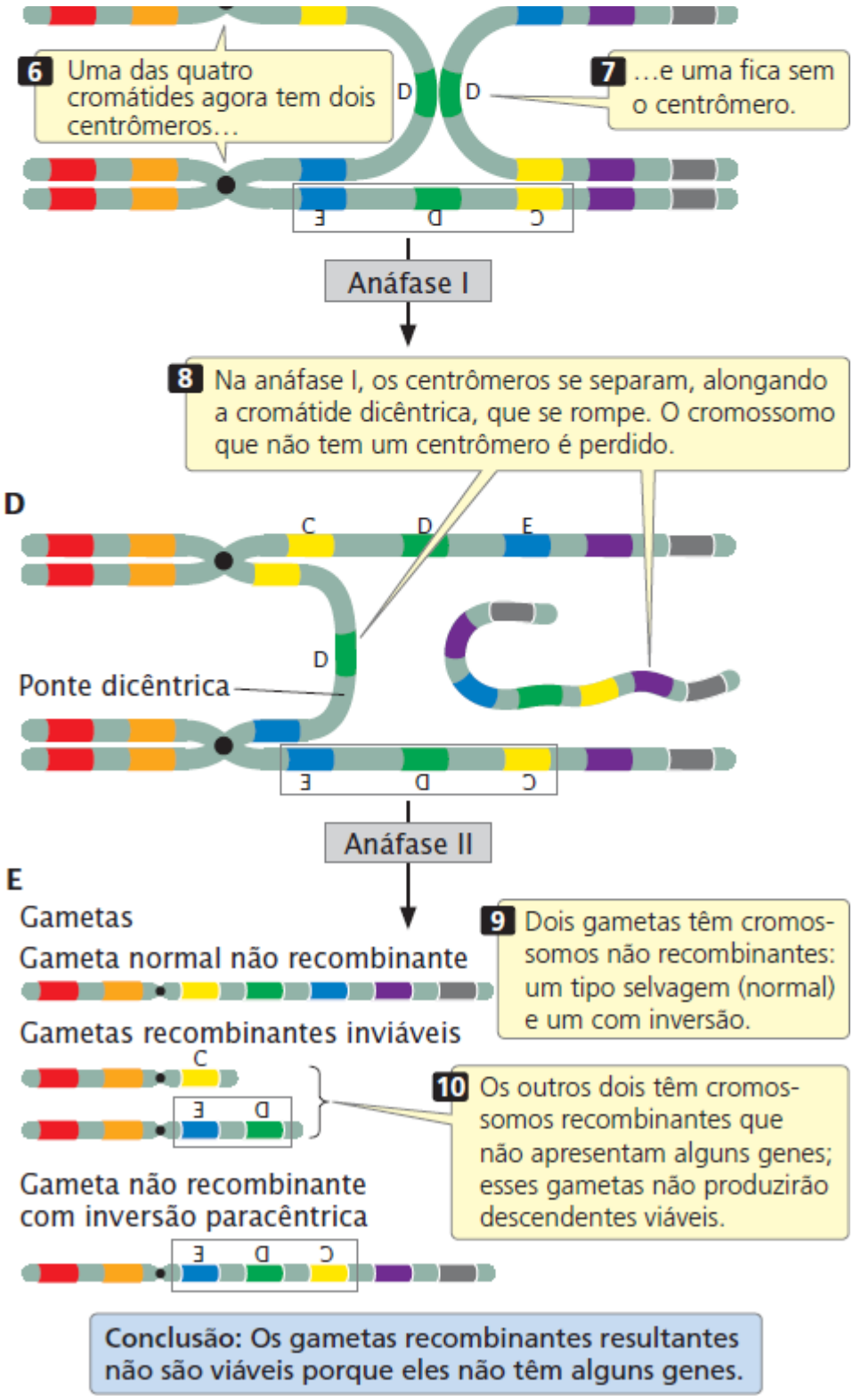


Figura 8.11 Em um indivíduo heterozigoto, um *crossing over* único em uma inversão paracêntrica gera gametas anormais.

Na segunda divisão da meiose, as cromátides-irmãs se separam e são produzidos quatro gametas (ver Figura 8.11 E). Dois deles têm os cromossomos originais, não recombinantes (AB•CDEFG e AB•EDC FG). Os outros dois apresentam cromossomos recombinantes que não têm alguns genes; esses gametas não produzem descendentes viáveis. Dessa forma, não surgem descendentes recombinantes quando o *crossing over* ocorre dentro de uma inversão paracêntrica. A chave é identificar que o *crossing over* ainda ocorre, mas, quando ocorre, os gametas recombinantes resultantes não são viáveis, então não são observados descendentes recombinantes.

A recombinação também é reduzida em uma inversão pericêntrica (**Figura 8.12**). Não são produzidas pontes dicêntricas nem fragmentos acêntricos, mas os cromossomos recombinantes têm cópias demais de alguns genes e nenhuma cópia de outros, então os gametas que recebem os cromossomos recombinantes não podem produzir descendentes viáveis.

As Figuras 8.11 e 8.12 ilustram os resultados dos *crossing overs* únicos dentro das inversões. Os *crossing overs* duplos nos quais ambos estão nas mesmas duas fitas (*crossing overs* duplos de fitas duplas) resultam em cromossomos recombinantes

funcionais (para ver por que os gametas funcionais são produzidos por *crossing overs* duplos, tente desenhar os resultados de um *crossing over* duplo de fitas duplas). Assim, mesmo que a taxa total de recombinação seja reduzida em uma inversão, alguns descendentes recombinantes viáveis ainda serão produzidos por meio dos *crossing overs* duplos de fitas duplas. **Resolva o Problema 27**

Importância das inversões na evolução. As inversões também podem ter um papel importante na evolução, ao suprimir a recombinação entre um conjunto de genes. Como já observamos, o *crossing over* em uma inversão de um indivíduo heterozigoto para uma inversão pericêntrica ou paracêntrica produz gametas não balanceados e nenhum descendente recombinante. Essa supressão da recombinação permite que conjuntos particulares de alelos coadaptados que funcionam bem juntos permaneçam intactos, não misturados pela recombinação.

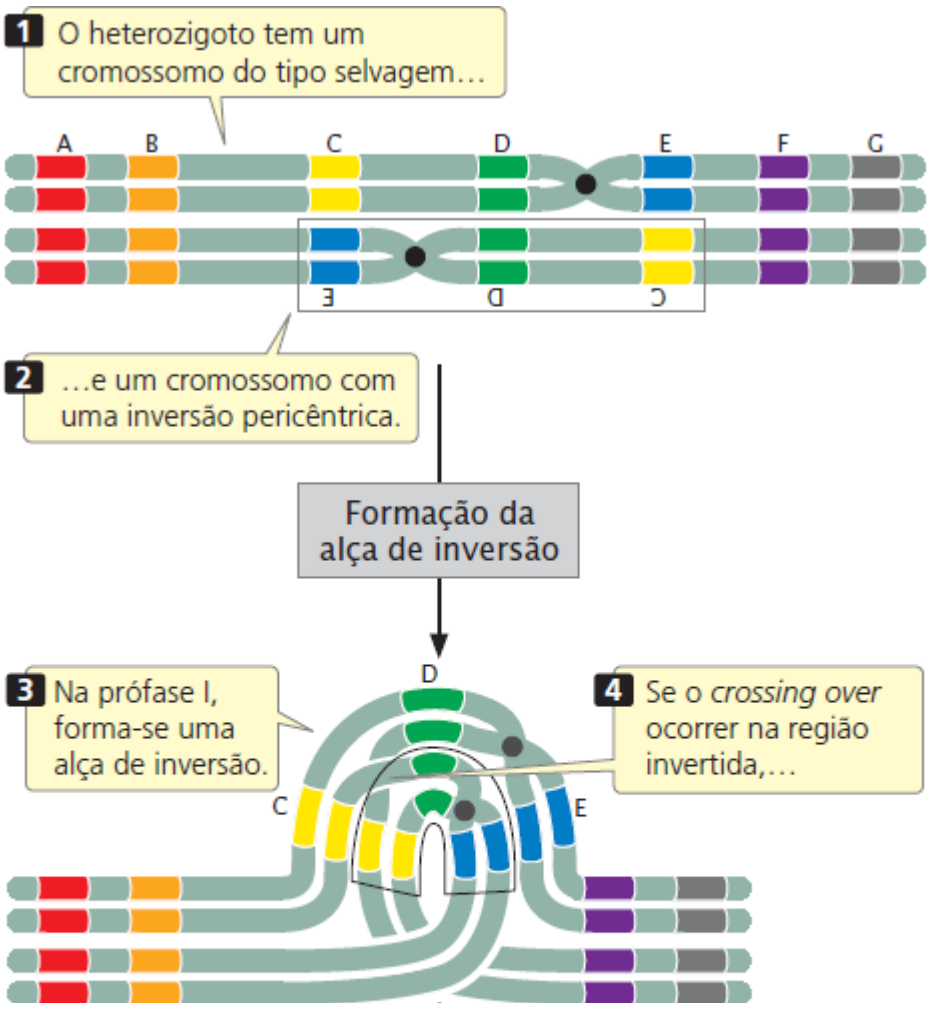
Conceitos

Em uma inversão, um segmento do cromossomo é movido 180º. As inversões provocam rupturas em alguns genes e podem deslocar outros genes para novos locais. Nos indivíduos heterozigotos para uma inversão cromossômica, os cromossomos homólogos formam uma alça na prófase I da meiose. Quando o *crossing over* ocorre na região invertida, geralmente são produzidos gametas inviáveis, resultando em redução das frequências de recombinação observadas.

Checagem dos conceitos 3

Um cromossomo dicêntrico é produzido quando o *crossing over* ocorre em um indivíduo heterozigoto para qual tipo de rearranjo cromossômico?

a. Duplicação.
b. Deleção.
c. Inversão paracêntrica.
d. Inversão pericêntrica.



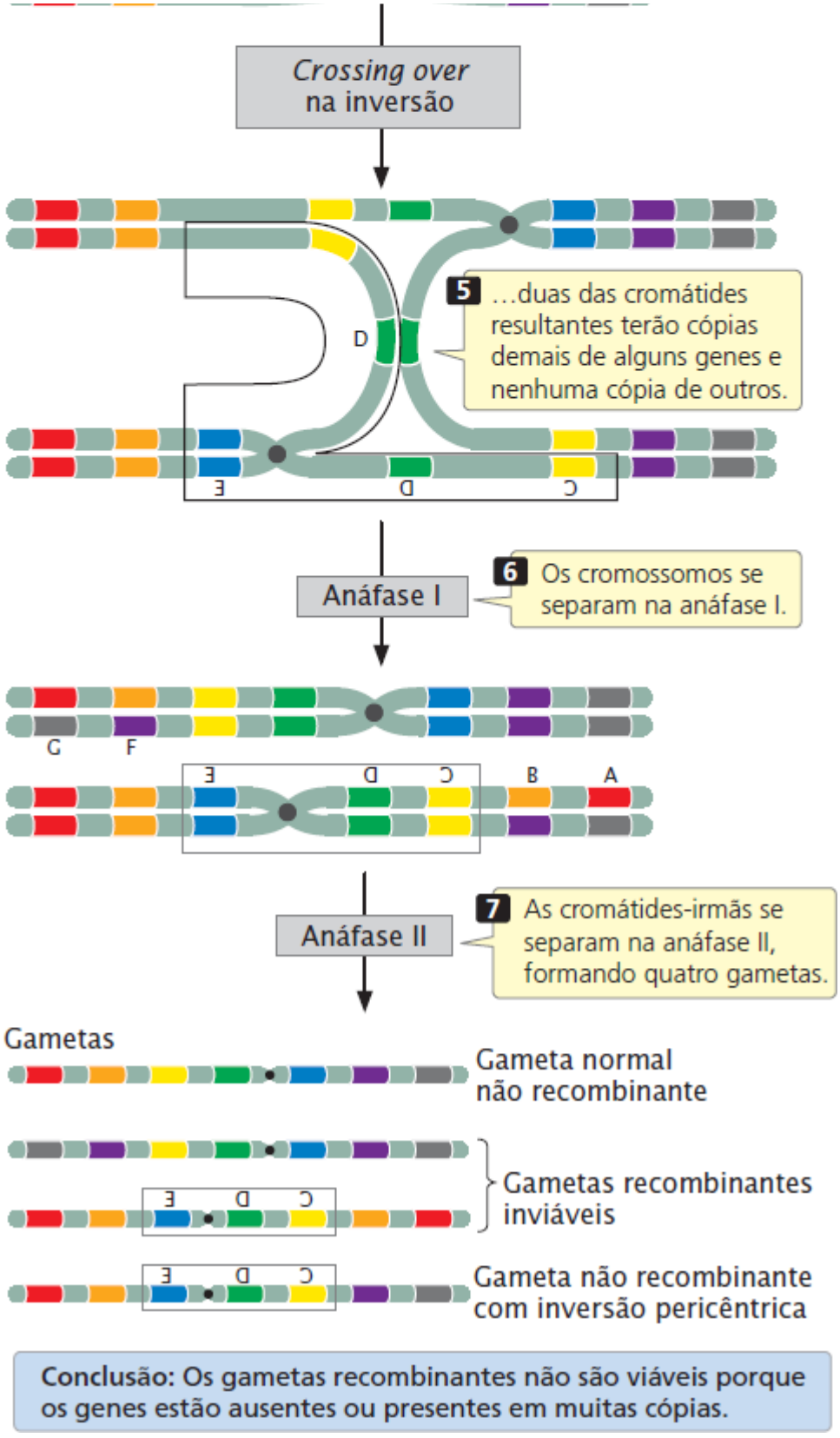


Figura 8.12 Em um indivíduo heterozigoto, um *crossing over* único em uma inversão pericêntrica gera gametas anormais.

Translocações

Uma **translocação** ocasiona o movimento do material genético entre cromossomos não homólogos (ver Figura 8.3 D) ou dentro do mesmo cromossomo. A translocação não deve ser confundida com *crossing over*, em que ocorre troca de material genético entre cromossomos *homólogos*.

Em uma **translocação não recíproca**, o material genético se desloca de um cromossomo para outro sem nenhuma troca recíproca. Considere os dois cromossomos não homólogos a seguir: AB•CDEFG e MN•OPQRS. Se o segmento EF se move do primeiro cromossomo para o segundo sem nenhuma transferência dos segmentos do segundo cromossomo para o primeiro, ocorre uma translocação não recíproca, produzindo os cromossomos AB•CDG e MN•OPEFQRS. Mais comumente, existe uma

troca de duas vias dos segmentos entre os cromossomos, resultando em uma **translocação recíproca**. Uma translocação recíproca entre os cromossomos AB•CDEFG e MN•OPQRS pode dar origem aos cromossomos AB•CDQRS e MN•OPEFG.

Efeitos das translocações. As translocações podem afetar um fenótipo de várias formas. Primeiro, elas podem unir fisicamente genes que antes estavam localizados em diferentes cromossomos. Essas novas uniões podem afetar a expressão do gene (um efeito da posição): os genes translocados para novos locais podem ficar sob o controle de diferentes sequências regulatórias ou de outros genes que afetam sua expressão.

Segundo, as rupturas de cromossomos que produzem as translocações podem ocorrer em um gene e comprometer sua função. Os geneticistas moleculares usam esses tipos de efeitos para mapear os genes humanos. A neurofibromatose é uma doença genética caracterizada por numerosos tumores fibrosos na pele e no tecido nervoso; ela é o resultado de uma mutação autossômica dominante. Estudos de ligação primeiro localizaram o *locus* que, quando mutado, provoca a neurofibromatose no cromossomo 17, mas a localização precisa do *locus* era desconhecida. Posteriormente, os geneticistas limitaram a localização quando identificaram dois pacientes com neurofibromatose que tinham uma translocação que afetava o cromossomo 17. Considera-se que esses pacientes desenvolveram neurofibromatose porque uma das rupturas de cromossomo que ocorreu na translocação atrapalhou um gene particular, resultando na doença. O DNA de regiões ao redor das rupturas foi sequenciado, levando à identificação do gene responsável pela neurofibromatose.

É frequente as deleções acompanharem as translocações. Em uma **translocação robertsoniana**, por exemplo, os braços longos de dois cromossomos acrocêntricos são unidos para um centrômero comum por uma translocação, gerando um cromossomo metacêntrico com dois braços longos e outro cromossomo com dois braços muito curtos (**Figura 8.13**). O cromossomo menor é perdido em muitos casos, porque os cromossomos muito pequenos não têm massa suficiente para se separar de forma adequada durante a mitose e meiose. O resultado é uma redução geral no número de cromossomos. Como veremos, as translocações robertsonianas são responsáveis por alguns casos da síndrome de Down, um distúrbio cromossômico discutido adiante neste capítulo.

Translocações na meiose. Os efeitos de uma translocação na segregação cromossômica durante a meiose dependem da natureza da translocação. Vamos considerar o que acontece em um indivíduo heterozigoto para uma translocação recíproca. Suponha que os cromossomos originais sejam AB•CDEFG e M•NOPQRST (indicados como N_1 e N_2, respectivamente, para os cromossomos 1 e 2 normais) e que uma translocação recíproca ocorre, produzindo cromossomos AB•CDQRS e M•NOPEFG (indicados como T_1 e T_2, respectivamente, para os cromossomos 1 e 2 translocados). Um indivíduo heterozigoto para essa translocação teria uma cópia normal de cada cromossomo e uma cópia translocada (**Figura 8.14 A**). Cada um desses cromossomos tem segmentos que são homólogos para outros dois cromossomos. Quando as sequências homólogas pareiam na prófase I da meiose, formam-se configurações semelhantes à cruz com todos os quatro cromossomos (**Figura 8.14 B**).

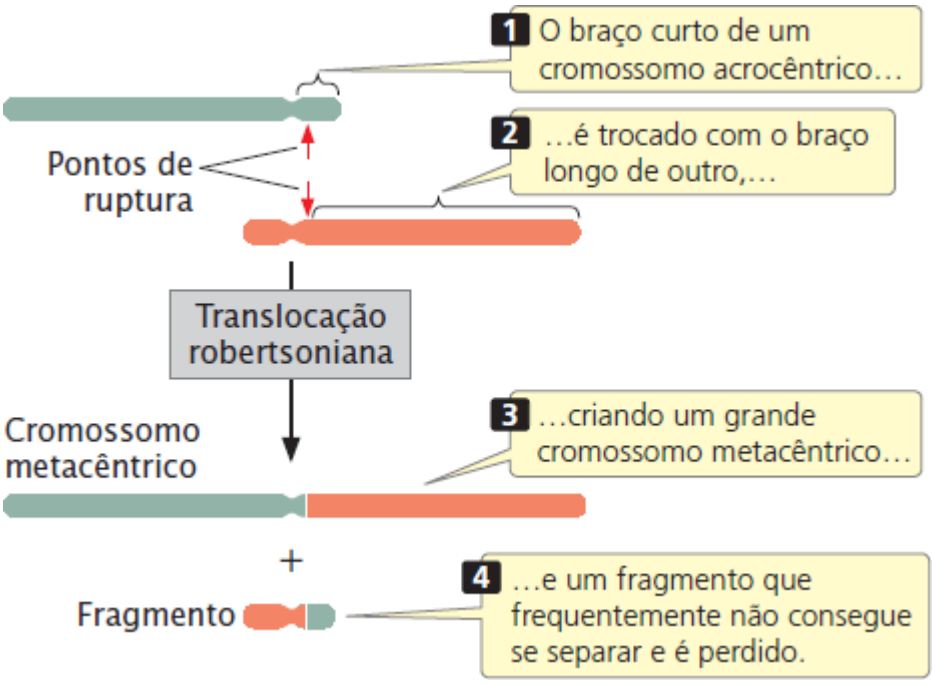


Figura 8.13 Em uma translocação robertsoniana, o braço curto de um cromossomo acrocêntrico é trocado com um braço longo de outro cromossomo.

Observe que N_1 e T_1 têm centrômeros homólogos (em ambos os cromossomos, o centrômero está entre os segmentos B e C);

da mesma forma, N_2 e T_2 têm centrômeros homólogos (entre os segmentos N e O). Normalmente, os centrômeros homólogos se separam e se movem na direção de polos opostos na anáfase I da meiose. Na translocação recíproca, os cromossomos podem se separar em três formas diferentes. Na **segregação alternada** (**Figura 8.14 C**), N_1 e N_2 se deslocam na direção de um polo, e T_1 e T_2 se deslocam para o polo oposto. Na **segregação adjacente-1**, N_1 e T_2 se deslocam na direção de um polo, e T_1 e N_2 se deslocam na direção do outro polo. Tanto na segregação alternada quanto na segregação adjacente-1, os centrômeros homólogos se separam no sentido de polos opostos. A **segregação adjacente-2**, na qual N_1 e T_1 se movem no sentido de um polo e T_2 e N_2 se movem no sentido do outro, é rara, porque em geral os dois cromossomos homólogos se separam na meiose.

Os produtos dos três padrões de separação estão ilustrados na **Figura 8.14 D**. Como você pode ver, os gametas produzidos pela separação alternada têm um conjunto completo de segmentos de cromossomo. Portanto, esses gametas são funcionais e podem produzir descendentes viáveis. Ao contrário, os gametas resultantes das segregações adjacente-1 e adjacente-2 não são viáveis porque alguns segmentos de cromossomos são encontrados em duas cópias, enquanto outros estão ausentes. Como a segregação adjacente-2 é rara, a maioria dos gametas é produzida pela segregação alternada ou adjacente-1. Portanto, espera-se que aproximadamente metade dos gametas de um indivíduo heterozigoto para uma translocação recíproca seja funcional.

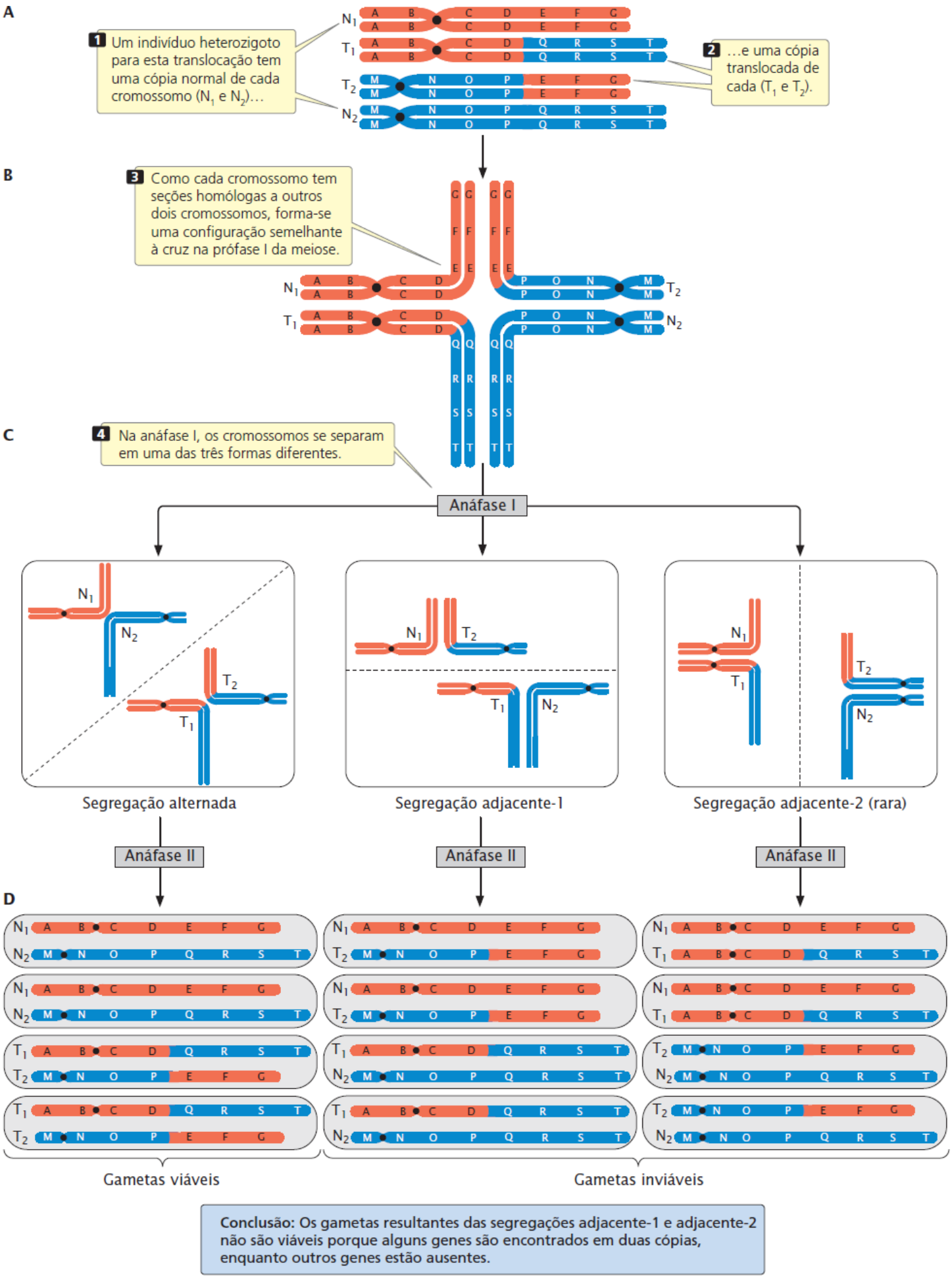


Figura 8.14 Em um indivíduo heterozigoto para uma translocação recíproca, as estruturas semelhantes à cruz formam um pareamento homólogo.

Importância das translocações na evolução. As translocações são, com frequência, importantes na evolução dos

cariótipos. Chimpanzés, gorilas e orangotangos têm 48 cromossomos, enquanto os seres humanos têm 46. O cromossomo 2 humano é um grande cromossomo metacêntrico, com padrões de bandeamento G que combinam com os encontrados em dois cromossomos acrocêntricos diferentes dos macacos (**Figura 8.15**). Aparentemente, uma translocação robertsoniana ocorreu em um ancestral humano, criando um grande cromossomo metacêntrico a partir de dois braços longos dos cromossomos acrocêntricos ancestrais, e um pequeno cromossomo que consiste em dois braços curtos. O pequeno cromossomo foi perdido, levando ao número reduzido de cromossomos nos seres humanos em relação aos cromossomos dos macacos. **Resolva o Problema 28**

Conceitos

Nas translocações, partes dos cromossomos se deslocam para outros cromossomos não homólogos ou para outras regiões do mesmo cromossomo. As translocações podem afetar o fenótipo ao fazer com que os genes se desloquem para novos locais, onde sofrem a influência de novas sequências reguladoras, ou ao quebrar os genes e prejudicar sua função.

Checagem dos conceitos 4

Qual é o desfecho de uma translocação robertsoniana?

a. Dois cromossomos acrocêntricos.
b. Um grande cromossomo metacêntrico e um cromossomo muito pequeno com dois braços muito curtos.
c. Um grande cromossomo metacêntrico e um grande cromossomo acrocêntrico.
d. Dois grandes cromossomos metacêntricos.

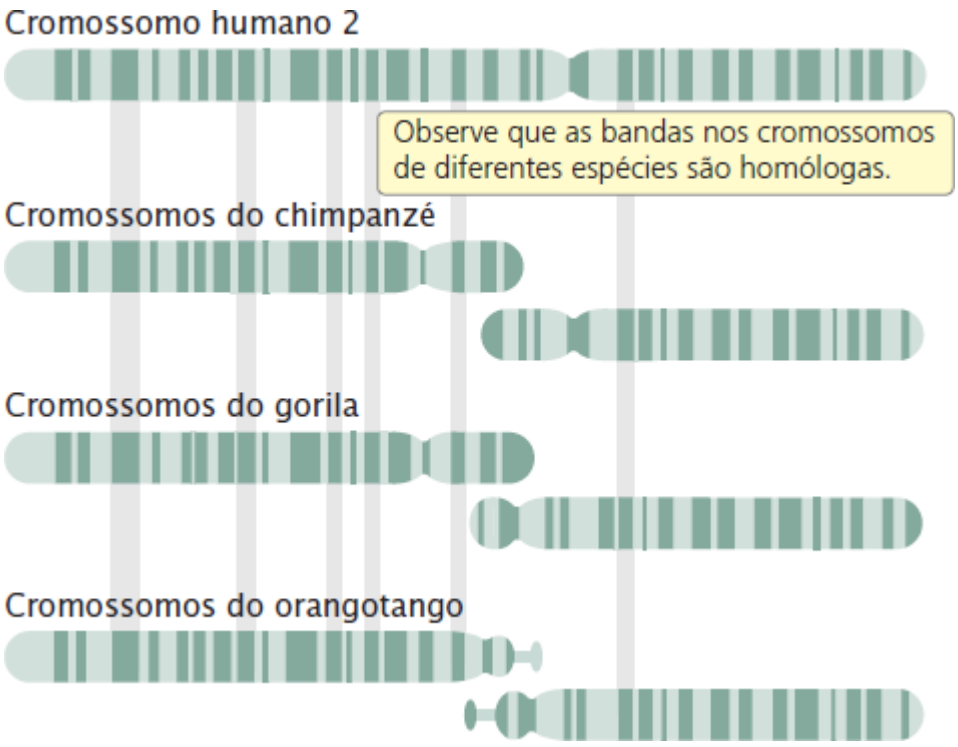


Figura 8.15 O cromossomo 2 humano tem uma translocação robertsoniana que não existe nos chimpanzés, gorilas ou orangotangos. O bandeamento G revela que uma translocação robertsoniana em um ancestral humano trocou os braços longos e curtos de dois cromossomos acrocêntricos que ainda são encontrados nos outros três primatas. Essa translocação criou o grande cromossomo 2 humano metacêntrico. As linhas conectoras em cinza destacam algumas, mas não todas as regiões da homologia entre os cromossomos.

Sítios frágeis

Os cromossomos das células que crescem em cultura às vezes desenvolvem constrições ou intervalos em locais chamados de **sítios frágeis** (**Figura 8.16**) porque são mais suscetíveis à ruptura em algumas condições. Mais de 100 sítios frágeis foram identificados nos cromossomos humanos.

Esses sítios são divididos em dois grupos. *Sítios frágeis comuns* são encontrados em todos os seres humanos e são uma característica comum dos cromossomos. Eles são a localização da ruptura e os rearranjos cromossômicos nas células do câncer,

levando a deleções cromossômicas, translocações e outros rearranjos cromossômicos. *Sítios frágeis raros* são encontrados em algumas pessoas e são herdados como um traço mendeliano. Os sítios frágeis raros estão associados a distúrbios genéticos, como déficit intelectual. A maioria deles são repetições expandidas de nucleotídios, nas quais o número de repetições de um conjunto de nucleotídios é aumentado (ver *Capítulo 18*).

Um dos sítios frágeis raros mais estudados está localizado no cromossomo X humano e é associado à **síndrome do X frágil**, um distúrbio que inclui déficit intelectual. Exibindo uma herança ligada ao X e surgindo com uma frequência de 1 em 5.000 meninos nascidos, foi demonstrado que a síndrome do X frágil resulta do aumento no número de repetições de um trinucleotídio CGG.

Estudos moleculares de sítios frágeis mostraram que muitos desses sítios têm mais de 100.000 pb de comprimento e incluem um ou mais genes. Frequentemente, os sítios frágeis são os últimos a serem replicados. Nesses sítios, as enzimas que replicam DNA podem atrasar enquanto a abertura do DNA continua (ver *Capítulo 12*), levando a longos trechos de DNA que estão abertos e vulneráveis à ruptura. Apesar dos avanços recentes na nossa compressão sobre os sítios frágeis, sua natureza ainda não é completamente compreendida.

Figura 8.16 Os sítios frágeis são regiões cromossômicas suscetíveis à ruptura em certas condições. Um local frágil no cromossomo X humano é apresentado aqui. (Cortesia de Dra. Christine Harrison.)

Variações no número de cópias

Os rearranjos cromossômicos são detectados ao examinar os cromossomos com um microscópio. O exame visual identifica os rearranjos cromossômicos com base nas mudanças no tamanho geral de um cromossomo, alteração dos padrões de bandeamento revelada pela coloração do cromossomo ou seu comportamento na meiose. O microscópio, entretanto, só

consegue detectar grandes rearranjos cromossômicos, tipicamente os que têm pelo menos 5 milhões de pares de base de comprimento.

Com o término do Projeto Genoma Humano (ver *Capítulo 20*), foram disponibilizadas informações sobre sequências de DNA encontradas em cromossomos individuais. Ao usar esta informação, os geneticistas conseguem agora examinar o número de cópias de sequências específicas de DNA presentes em uma célula e detectar duplicações, deleções e outros rearranjos cromossômicos que não podem ser observados apenas ao microscópio. O trabalho foi muito facilitado com a disponibilidade de *microarrays* (ver *Capítulo 20*), que permitem a detecção simultânea de centenas de milhares de sequências específicas de DNA pelo genoma. Como estes métodos medem o número de cópias de sequências específicas de DNA, as variações que eles detectam são chamadas de **variações no número de cópias** (CNVs). Elas incluem duplicações e deleções que variam de milhares a milhões de pares de bases. Muitas destas variações incluem pelo menos um gene e podem incluir vários genes.

Estudos recentes da CNV revelaram que as duplicações e deleções cromossômicas submicroscópicas são muito comuns: a pesquisa sugere que cada pessoa pode ter cerca de 1.000 CNVs. É provável que muitas não tenham efeitos fenotípicos observáveis, mas algumas CNVs estão envolvidas em várias doenças e distúrbios. Por exemplo, Janine Wagenstaller *et al*. estudaram a CNV em 67 crianças com déficit intelectual sem explicação e descobriram que 11 (16%) delas apresentavam duplicações ou deleções. Essas variações também estão associadas a osteoporose, autismo, esquizofrenia e várias outras doenças e distúrbios. Resolva o Problema 20

Conceitos

As variações no número de cópias de sequências particulares de DNA (variações no número de cópias) são muito comuns no genoma humano.

8.3 Aneuploidia é o aumento ou a redução no número de cromossomos individuais

Além dos rearranjos cromossômicos, as mutações cromossômicas incluem alterações do número de cromossomos. As variações do número de cromossomos podem ser classificadas em dois tipos básicos: **aneuploidia**, que é uma alteração no número de cromossomos, e **poliploidia**, uma alteração do número de conjuntos de cromossomos.

A aneuploidia pode surgir de várias formas. Primeiro, um cromossomo pode ser perdido durante a mitose ou meiose, se, por exemplo, seu centrômero for eliminado. A perda do centrômero impede que as fibras do fuso se prendam; então, o cromossomo não se move para o polo do fuso e não é incorporado ao núcleo após a divisão celular. Segundo, o pequeno cromossomo gerado por uma translocação robertsoniana pode ser perdido na mitose ou meiose. Terceiro, as aneuploidias podem surgir pela não disjunção, a falha dos cromossomos homólogos ou cromátides-irmãs em se separar na meiose ou mitose (ver *Capítulo 4*). A não disjunção faz com que alguns gametas ou células tenham um cromossomo extra e outros gametas e células não tenham um cromossomo (**Figura 8.17**). Resolva o Problema 29

Tipos de aneuploidia

Vamos considerar quatro tipos de aneuploidias em indivíduos diploides: nulissomia, monossomia, trissomia e tetrassomia.

1. A **nulissomia** é a perda de ambos os membros de um par homólogo de cromossomos. É representada como $2n - 2$, em que n se refere ao número haploide de cromossomos. Assim, nos seres humanos, que normalmente têm $2n = 46$ cromossomos, um zigoto nulissômico tem 44 cromossomos.
2. A **monossomia** é a perda de um único cromossomo, representada como $2n - 1$. Um zigoto humano monossômico tem 45 cromossomos.
3. A **trissomia** é o ganho de um único cromossomo, representada como $2\ n + 1$. Um zigoto humano trissômico tem 47 cromossomos. O ganho de um cromossomo significa que existem três cópias homólogas de um cromossomo. A maioria dos casos da síndrome de Down, discutidos adiante neste capítulo, é o resultado da trissomia do cromossomo 21.
4. A **tetrassomia** é o ganho de dois cromossomos homólogos, representada como $2n + 2$. Um zigoto humano tetrassômico tem 48 cromossomos. A tetrassomia *não* é o ganho de dois cromossomos extras, mas, em vez do ganho de dois cromossomos homólogos, teremos quatro cópias homólogas de um dado cromossomo.

Mais de uma mutação aneuploide pode ocorrer no mesmo organismo. Um indivíduo que tem uma cópia extra de dois cromossomos diferentes (não homólogos) é chamado de trissômico duplo e é representado como $2n + 1 + 1$. Da mesma forma, um monossômico duplo tem dois cromossomos não homólogos a menos ($2n - 1 - 1$), e um tetrassômico duplo tem dois pares

extras de cromossomos homólogos ($2n + 2 + 2$).

Efeitos da aneuploidia

Um dos primeiros aneuploides a ser identificado foi uma mosca-da-fruta com um único cromossomo X e nenhum cromossomo Y, descoberta por Calvin Bridges em 1913 (ver *Capítulo 4*). Outro estudo inicial sobre aneuploidia se concentrava nos mutantes no estramônio, *Datura stramonium*. A. Francis Blakeslee começou a criar essa planta em 1913, e observou que os cruzamentos com vários mutantes do estramônio produziam razões incomuns de descendentes. Por exemplo, o mutante *globo* (que produzia uma casca de semente globosa) era dominante, mas herdado principalmente da genitora. Quando as plantas com a mutação *globo* eram autofertilizadas, apenas 25% dos descendentes tinham o fenótipo globoso. Se o mutante *globo* era rigorosamente dominante, Blakeslee devia ter observado 75% dos descendentes com o traço (ver *Capítulo 3*), e então os 25% que ele observou eram incomuns. Blakeslee isolou 12 mutantes diferentes (**Figura 8.18**) que exibiam padrões peculiares de herança. Posteriormente, John Belling demonstrou que essas 12 mutações eram, na verdade, trissomias. *Datura stramonium* tem 12 pares de cromossomos ($2n = 24$), e cada um dos 12 mutantes é trissômico para um diferente par de cromossomos. A natureza aneuploide dos mutantes explica as razões incomuns que Blakeslee observou nos descendentes. Muitos dos cromossomos extras nos trissômicos são perdidos na meiose, então menos de 50% dos gametas carreiam cromossomos extras, e a proporção de trissômicos nos descendentes foi baixa. Além disso, o pólen com um cromossomo extra não foi tão bem-sucedido na fertilização, e os zigotos trissômicos eram menos viáveis.

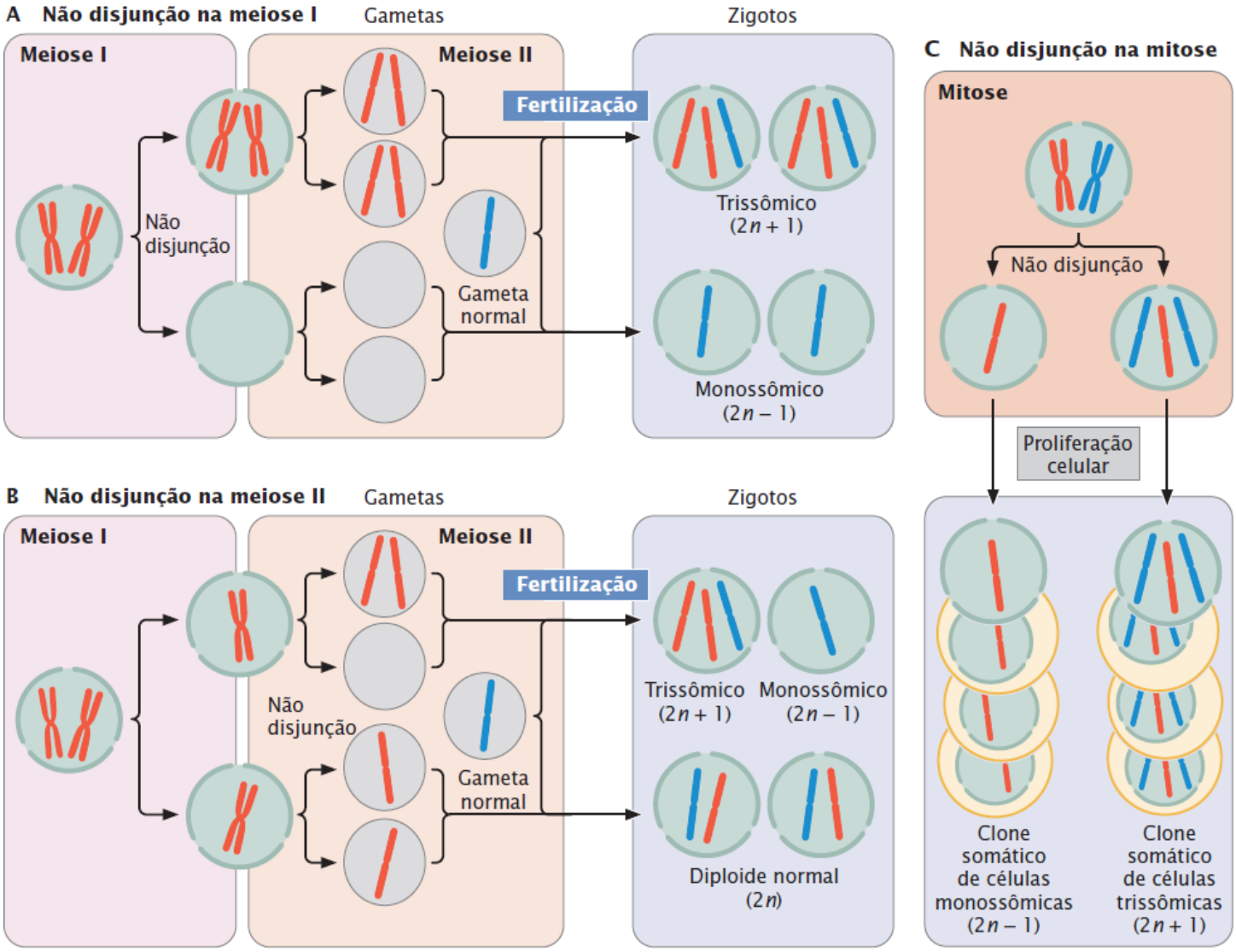


Figura 8.17 Os aneuploides podem ser produzidos pela não disjunção na meiose I, na meiose II e na mitose. Os gametas resultantes da meiose com não disjunção se combinam com um gameta (com o cromossomo azul) resultante da meiose normal para produzir os zigotos.

Em geral, a aneuploidia altera o fenótipo de forma drástica. Na maioria dos animais e em muitas plantas, as mutações aneuploides são letais. Como a aneuploidia afeta o número de cópias de genes, mas não sua sequência de nucleotídios, é mais provável que os efeitos da aneuploidia se devam à dosagem genética anormal. Ela altera a dosagem para alguns, mas não todos,

genes, alterando as concentrações relativas dos produtos dos genes e interferindo no desenvolvimento normal.

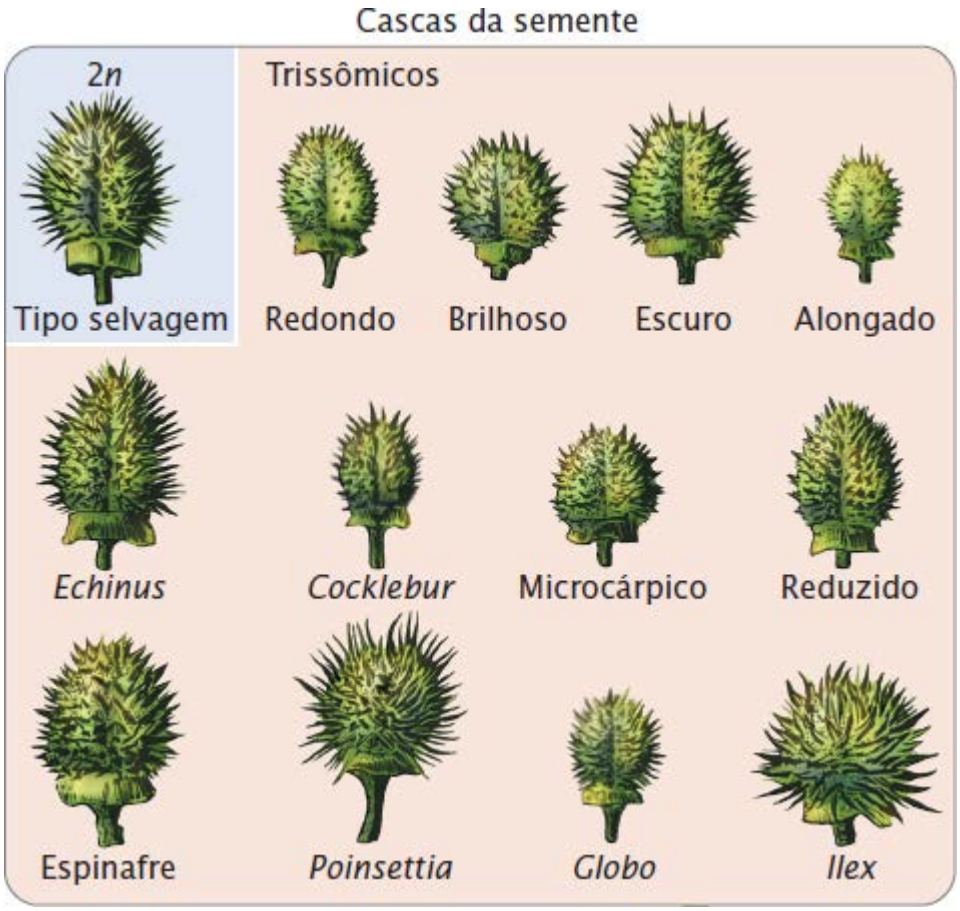


Figura 8.18 As cascas de sementes mutantes na trombeta (*Datura stramonium*) são o resultado de diferentes trissomias. Cada tipo de casca de semente é um fenótipo trissômico para um diferente cromossomo.

Uma exceção importante para a relação entre o número de genes e a dosagem genética são os genes no cromossomo X dos mamíferos. Nesses animais, a inativação do cromossomo X garante que os machos (que têm um único cromossomo X) e as fêmeas (que têm dois cromossomos X) recebam a mesma dosagem funcional para os genes ligados ao X (ver *Capítulo 4* para uma discussão sobre a inativação do cromossomo X). Os cromossomos X adicionais são inativados nos mamíferos, então podemos esperar que a aneuploidia dos cromossomos sexuais seja menos prejudicial nesses animais. De fato, esse é o caso para camundongos e seres humanos, para os quais os aneuploides dos cromossomos sexuais são a forma mais comum de aneuploidia observada nos organismos vivos. É provável que os aneuploides do cromossomo Y sejam comuns porque existe muito pouca informação nesse cromossomo.

Conceitos

A aneuploidia, a perda ou o ganho de um ou mais cromossomos, pode surgir a partir da perda de um cromossomo após a translocação ou não disjunção na meiose ou mitose. Ela atrapalha a dosagem genética e, em muitos casos, tem graves efeitos fenotípicos.

Checagem dos conceitos 5

Um organismo diploide tem $2n = 36$ cromossomos. Quantos cromossomos serão encontrados em um membro trissômico dessa espécie?

Aneuploidia nos seres humanos

Por motivos desconhecidos, uma elevada porcentagem de todos os embriões humanos concebidos tem anomalias cromossômicas. Os achados de estudos de mulheres que tentavam engravidar sugerem que mais de 30% das concepções sofrem aborto espontâneo, em geral no início do desenvolvimento, quando a mulher não tem ideia de que esteja grávida. Defeitos cromossômicos são encontrados em pelo menos 50% dos fetos humanos abortados de forma espontânea, e a aneuploidia é responsável pela maioria dos casos. Esta taxa de anomalia cromossômica nos seres humanos é maior que nos outros organismos

estudados; nos camundongos, por exemplo, a aneuploidia é encontrada em no máximo 2% dos óvulos fertilizados. A aneuploidia nos seres humanos produz graves problemas de desenvolvimento que levam ao aborto espontâneo. Apenas cerca de 2% dos fetos com um defeito cromossômico sobrevivem ao nascimento.

Aneuploides dos cromossomos sexuais. As aneuploidias mais comuns observadas nos seres humanos vivos são as que envolvem os cromossomos sexuais. Como ocorre com todos os mamíferos, a aneuploidia dos cromossomos sexuais humanos é mais bem tolerada do que dos cromossomos autossômicos. As síndromes de Turner e Klinefelter (ver Figuras 4.8 e 4.9) são resultados de aneuploidia dos cromossomos sexuais.

Aneuploides autossômicos. Os aneuploides autossômicos que resultam em nascidos vivos são menos comuns do que os aneuploides dos cromossomos sexuais nos seres humanos, provavelmente porque não há mecanismo de compensação de dosagem para os cromossomos autossômicos. A maioria dos aneuploides autossômicos é abortada de forma espontânea, apesar de os aneuploides de alguns dos autossomos pequenos, como o cromossomo 21, completarem o desenvolvimento e resultarem em uma pessoa com aneuploidia. Como esses cromossomos são menores e carreiam menos genes, a existência de cópias extras é menos prejudicial do que nos cromossomos maiores.

Síndrome de Down. Em 1866, John Langdon Down, médico e superintendente do Earslwood Asylum em Surrey, Inglaterra, observou uma incrível semelhança entre vários de seus pacientes intelectualmente incapacitados: todos apresentavam rosto plano, redondo, nariz pequeno e olhos ovais. Suas características eram tão semelhantes, na verdade, que ele sentia como se fossem facilmente confundidos como crianças da mesma família. Down não compreendia a causa do déficit intelectual de seus pacientes, mas sua descrição original registrava de modo fiel as características físicas dessa forma genética de déficit intelectual. Em sua homenagem, esse distúrbio é conhecido como síndrome de Down.

A **síndrome de Down**, também conhecida como **trissomia do 21**, é a aneuploidia autossômica mais comum nos seres humanos (**Figura 8.19 A**). A incidência da síndrome de Down nos EUA é semelhante à observada no mundo – cerca de 1 em 700 nascimentos –, embora a incidência aumente entre as crianças de mães mais velhas. Aproximadamente 92% das pessoas com síndrome de Down têm três cópias completas do cromossomo 21 (e, portanto, um total de 47 cromossomos), uma condição chamada de **síndrome de Down primária** (**Figura 8.19 B**). A síndrome de Down primária geralmente surge a partir da não disjunção espontânea na formação do óvulo: cerca de 75% dos eventos de não disjunção que provocam a síndrome de Down são de origem materna, a maioria surgindo na meiose I. A maior parte das crianças com síndrome de Down são filhos de pais normais, e a falha dos cromossomos em se dividirem tem pouca tendência hereditária. Um casal que tenha um filho com a síndrome de Down primária tem um risco um pouco maior de ter um segundo filho com a síndrome (em comparação a outros casais de idade semelhante que não têm nenhum filho com síndrome de Down). Da mesma forma, não é mais provável que os parentes do casal tenham um filho com a síndrome de Down primária.

Cerca de 4% das pessoas com síndrome de Down não são trissômicas para um cromossomo 21 completo. Em vez disso, eles têm 46 cromossomos, mas uma cópia extra da parte do cromossomo 21 é presa a outro cromossomo por meio de uma translocação. Essa condição é chamada de **síndrome de Down familiar**, porque tem uma tendência de surgir nas famílias. As características fenotípicas da síndrome de Down familiar são as mesmas que as da síndrome de Down primária.

A síndrome de Down familiar surge nos descendentes cujos pais são portadores de cromossomos que sofreram uma translocação robertsoniana, mais comumente entre o cromossomo 21 e o cromossomo 14; o braço longo do 21 e o braço curto do 14 trocam de lugar (**Figura 8.20**). Essa troca produz um cromossomo que inclui os braços longos dos cromossomos 14 e 21 e um cromossomo muito pequeno com os braços curtos dos cromossomos 21 e 14. O pequeno cromossomo é perdido após várias divisões celulares. Embora a troca entre os cromossomos 21 e 14 seja a causa mais comum da síndrome de Down familiar, a condição também pode ser causada por translocações entre o cromossomo 21 e outros cromossomos, como o 15.

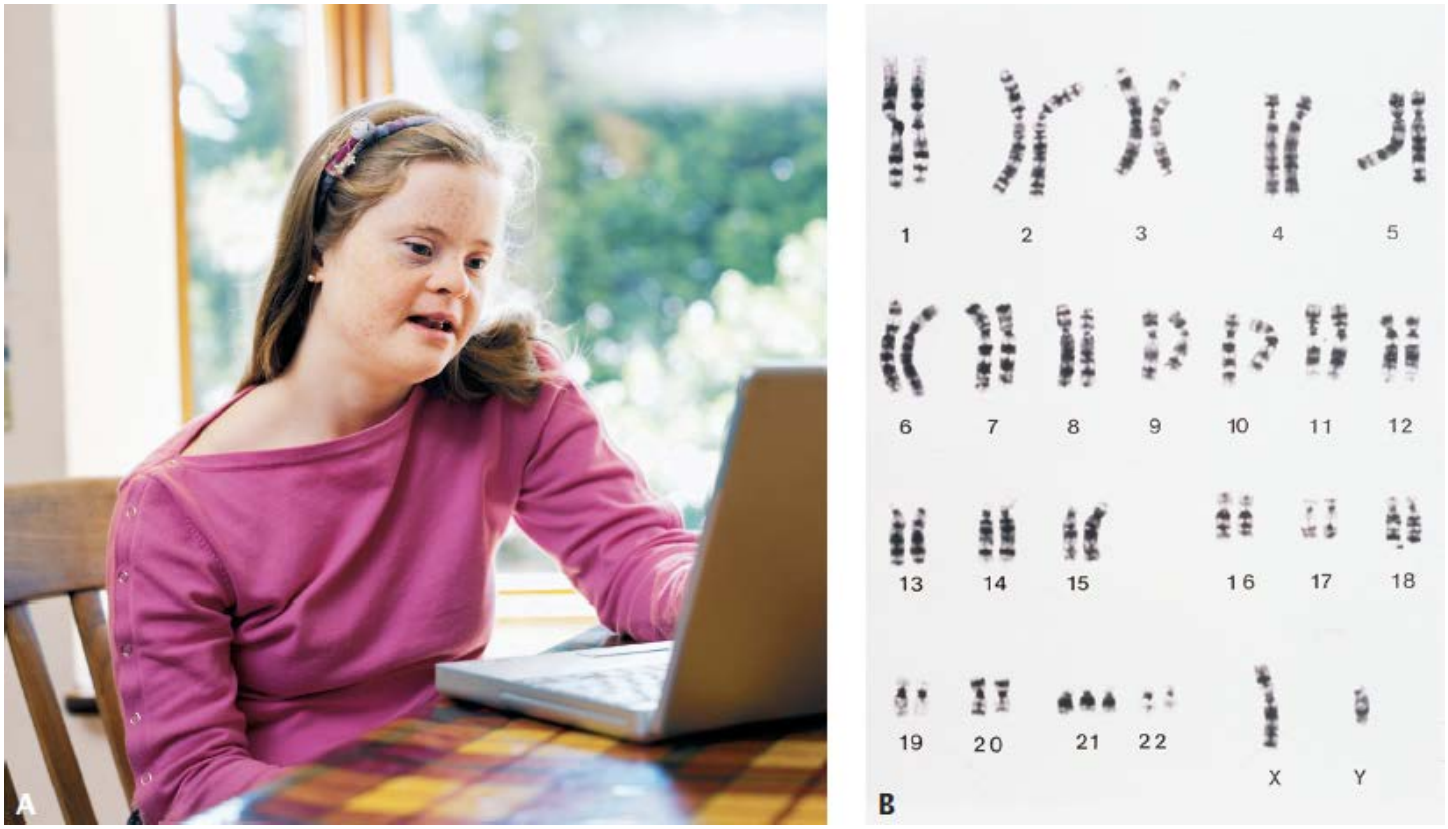


Figura 8.19 A síndrome de Down (A) é causada por uma trissomia do cromossomo 21 (B). (*Parte A*: George Doyle/Stockbyte/Getty Images. *Parte B*: L. Wilatt, East Anglian Regional Genetics Service/Science Photo Library/Photo Researchers.)

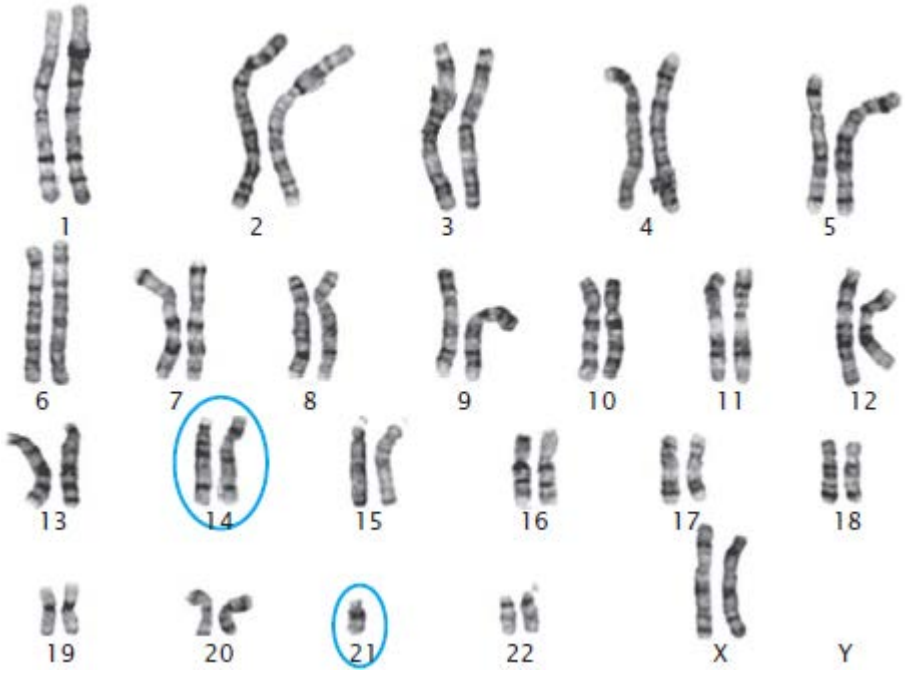


Figura 8.20 A translocação do cromossomo 21 para outro cromossomo resulta na síndrome de Down. Aqui, o braço longo do cromossomo 21 está ligado ao cromossomo 14. Esse cariótipo é de um portador da translocação, que é normal do ponto de vista do fenótipo, mas com maior risco de ter filhos com a síndrome de Down. (© Centre for Genetics Education para e em nome de Crown no direito do Estado de New South Wales, Australia.)

As pessoas com a translocação, chamadas de **portadores da translocação**, não têm a síndrome de Down. Embora eles apresentem apenas 45 cromossomos, seus fenótipos são normais porque eles têm duas cópias dos braços longos dos cromossomos 14 e 21 e aparentemente, os braços curtos desses cromossomos (que são perdidos) não carreiam informações genéticas essenciais. Apesar de os portadores da translocação serem completamente saudáveis, eles apresentam maior chance de ter crianças com a síndrome de Down.

Quando esse portador produz os gametas, o cromossomo da translocação pode se separar de três formas diferentes. Primeiro, ele pode se separar a partir dos cromossomos 14 e 21 normais na anáfase I da meiose (**Figura 8.21 A**). Nesse tipo de separação,

metade dos gametas terá o cromossomo da translocação e nenhuma outra cópia dos cromossomos 21 e 14; a fusão desse gameta com um gameta normal dará origem a um portador da translocação. A outra metade dos gametas produzidos por esse tipo de se separação será normal, cada um com uma única cópia dos cromossomos 21 e 14, e resultará em descendentes normais.

Por outro lado, o cromossomo da translocação pode se separar do cromossomo 14 e passar para a mesma célula com o cromossomo 21 normal (**Figura 8.21 B**). Esse tipo de separação produz apenas gametas anormais; metade terá duas cópias funcionais do cromossomo 21 (uma normal e a outra presa ao cromossomo 14) e a outra metade não terá o cromossomo 21. Se um gameta com as duas cópias funcionais do cromossomo 21 se une a um gameta normal que carreia uma única cópia desse cromossomo, o zigoto resultante terá síndrome de Down familiar. Se o gameta sem o cromossomo 21 se unir a um gameta normal, o zigoto resultante terá a monossomia do 21 e será abortado espontaneamente.

No terceiro tipo de separação, o cromossomo da translocação e a cópia normal do cromossomo 14 se separam juntos (**Figura 8.21 C**). Esse padrão é raro, porque os dois centrômeros são derivados do cromossomo 14 e, em geral, se separam um do outro. Todos os gametas produzidos por esse processo serão anormais: metade resulta na monossomia do 14 e a outra metade na trissomia do 14. Todos são espontaneamente abortados.

Assim, apenas três dos seis tipos de gametas que podem ser produzidos por um portador de translocação resultarão no nascimento de um feto e, teoricamente, esses gametas devem surgir na mesma frequência. Um terço dos descendentes de um portador da translocação deve ser portador como seu pai/mãe; um terço deve ter síndrome de Down familiar; e um terço deve ser normal. Na verdade, entretanto, menos de um terço dos filhos nascidos de portadores da translocação tem a síndrome de Down, o que sugere que alguns dos embriões com a síndrome de Down são espontaneamente abortados. Resolva o Problema 32

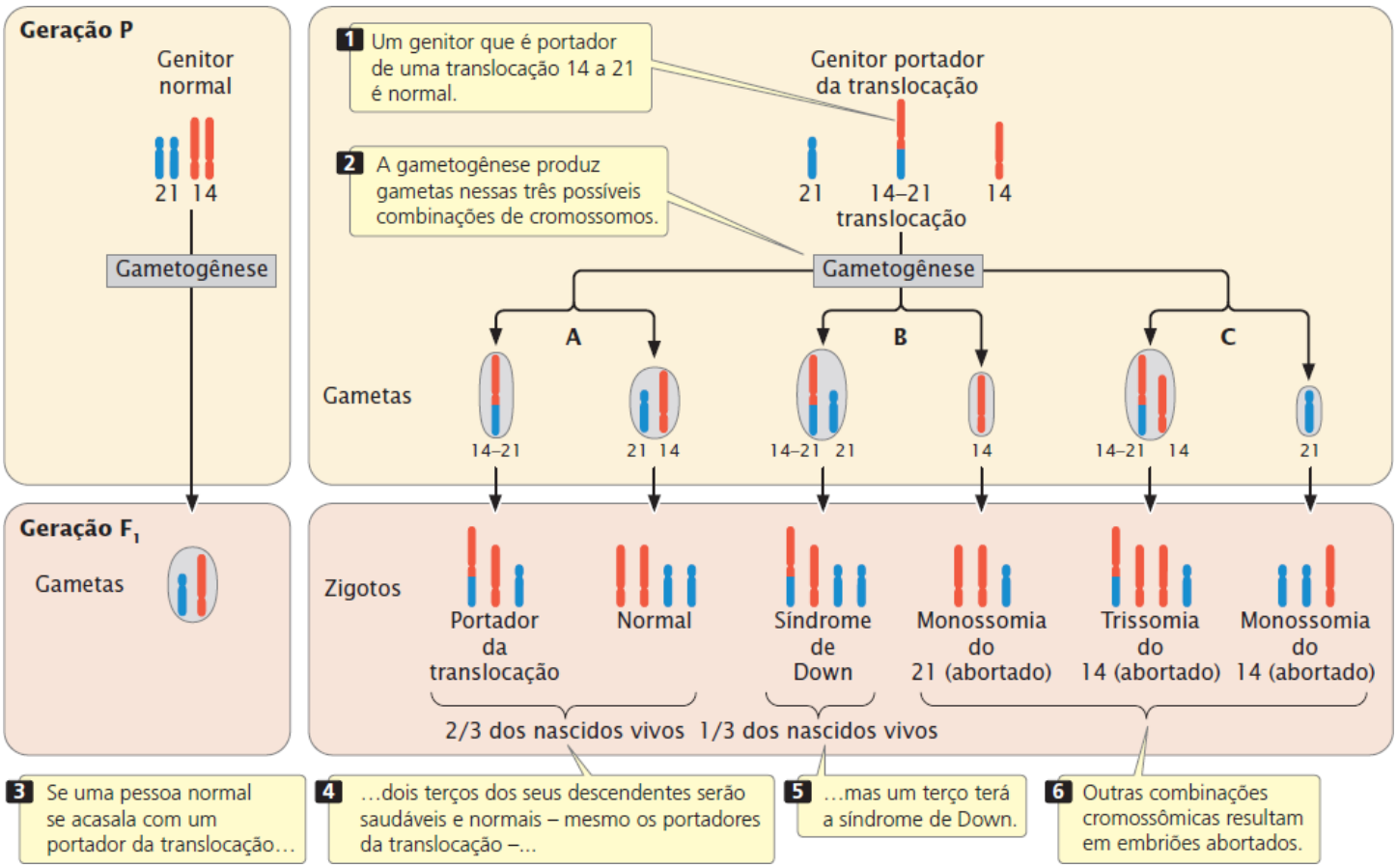


Figura 8.21 Os portadores da translocação correm maior risco de terem filhos com a síndrome de Down.

Outras trissomias humanas. Poucos humanos aneuploides autossômicos além da trissomia do 21 conseguem sobreviver. A **trissomia do 18**, também conhecida como **síndrome de Edward**, tem uma frequência de aproximadamente 1 em 8.000 nascidos vivos. Os recém-nascidos com essa síndrome têm grave déficit intelectual, orelhas de implantação baixa, pescoço curto, pés deformados, dedos das mãos cerrados, problemas cardíacos e outras incapacidades. Poucos vivem por mais de 1 ano. A **trissomia do 13** tem uma frequência de cerca de 1 em 15.000 nascidos vivos e produz características que são coletivamente conhecidas com **síndrome de Patau**. As características dessa condição incluem grave déficit intelectual, cabeça pequena, testa inclinada, olhos pequenos, fenda labiopalatina, dedos extras nas mãos e nos pés e vários outros problemas. Cerca de metade das crianças com trissomia do 13 morre no primeiro mês de vida, e 95% morrem até os 3 anos de idade. A **trissomia do 8** é ainda

mais rara, e surge com uma frequência que varia de 1 em 25.000 a 1 em 50.000 nascidos vivos. Esse aneuploide é caracterizado por déficit intelectual, dedos das mãos e pés contraídos, orelhas de implantação baixa e malformadas e testa proeminente. Muitos com essa condição têm uma expectativa de vida normal.

Aneuploidia e idade materna. A maioria dos casos de síndrome de Down e outros tipos de aneuploidia nos seres humanos surgem a partir de não disjunção materna, e a frequência de aneuploidia aumenta com a idade materna (**Figura 8.22**). Não se sabe por que a idade materna está associada à não disjunção. As fêmeas dos mamíferos nascem com os ovócitos primários suspensos no subestágio diplóteno da prófase I da meiose. Pouco antes da ovulação, a meiose termina e a primeira divisão é completa, produzindo um ovócito secundário. Nesse ponto, a meiose é suspensa novamente e permanece assim até que o ovócito secundário seja penetrado por um espermatozoide. A segunda divisão meiótica ocorre imediatamente antes de os núcleos do óvulo e do espermatozoide se unirem para formar um zigoto.

Os ovócitos primários permanecem suspensos no diplóteno por muitos anos antes de a ovulação ocorrer e a meiose recomeçar. Os componentes do fuso e outras estruturas necessárias para a separação cromossômica podem se romper na longa interrupção da meiose, gerando mais aneuploidia nos filhos das mulheres mais velhas. De acordo com essa teoria, não é observado o efeito da idade nos homens, porque os espermatozoides são produzidos de forma contínua após a puberdade, sem suspensão prolongada das divisões meióticas.

Aneuploidia e câncer. Muitas células tumorais têm cromossomos extras ou ausentes, ou ambos; alguns tipos de tumores estão associados a mutações cromossômicas específicas, incluindo aneuploidia e rearranjos cromossômicos. O papel das mutações cromossômicas no câncer será explorado no Capítulo 23.

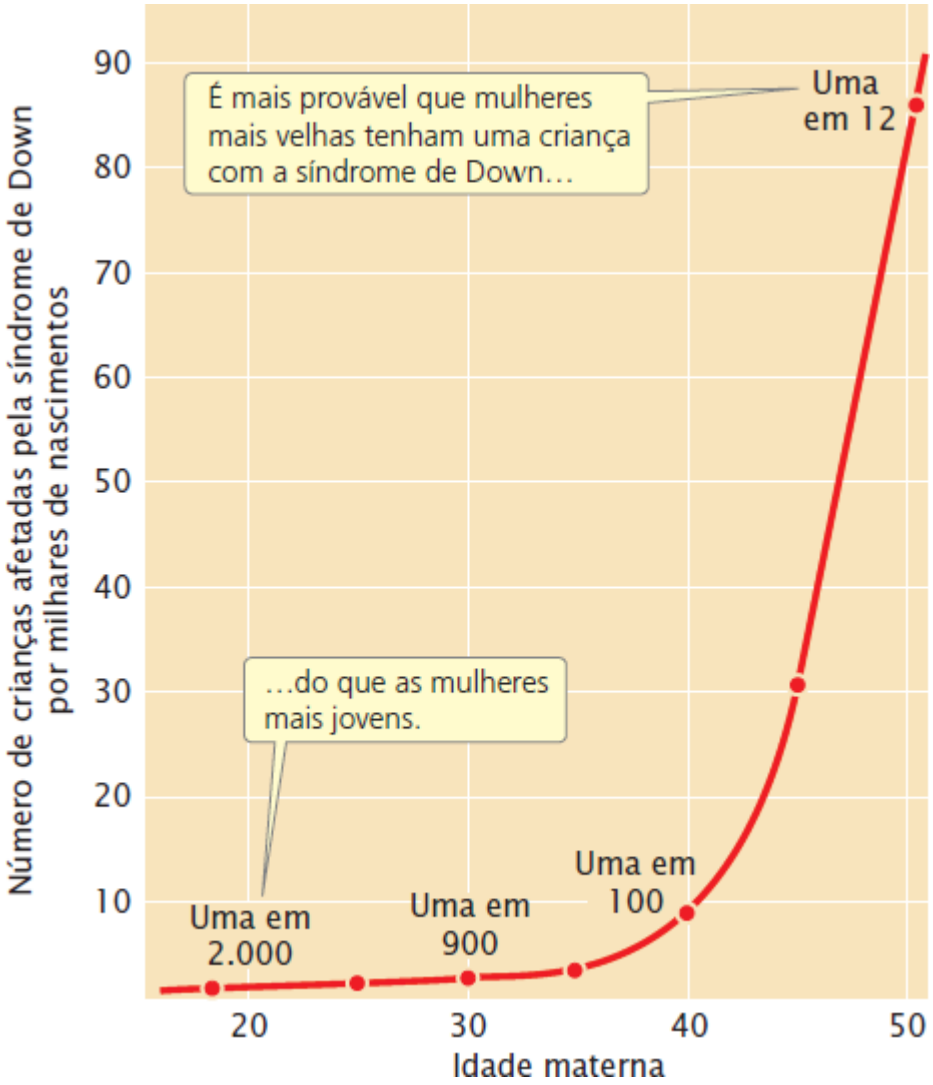


Figura 8.22 **A incidência da síndrome de Down primária e de outros aneuploides aumenta com a idade materna.**

Conceitos

Nos seres humanos, as aneuploidias dos cromossomos sexuais são mais comuns do que as aneuploidias autossômicas. A inativação do cromossomo X evita problemas da dosagem gênica para os genes ligados ao

X. A síndrome de Down é o resultado de três cópias funcionais do cromossomo 21, mesmo por meio da trissomia (síndrome de Down primária) ou de uma translocação robertsoniana (síndrome de Down familiar).

Checagem dos conceitos 6

Explique resumidamente por que nos seres humanos e nos mamíferos os aneuploides dos cromossomos sexuais são mais comuns do que os aneuploides autossômicos.

Dissomia uniparental

Normalmente, os cromossomos de um par homólogo são herdados de genitores diferentes – um do pai e outro da mãe. O desenvolvimento de técnicas moleculares que facilitam a identificação de sequências específicas de DNA (ver *Capítulo 19*) tornou possível a determinação da origem paterna/materna dos cromossomos. De modo surpreendente, algumas vezes ambos os cromossomos são herdados do mesmo genitor, uma condição chamada de **dissomia uniparental**.

É provável que muitos casos de dissomia uniparental se originem como uma trissomia. Embora a maioria das trissomias autossômicas seja letal, um embrião trissômico consegue sobreviver se um dos três cromossomos for perdido no início do desenvolvimento. Se, por acaso, os dois cromossomos remanescentes forem do mesmo genitor, surge a dissomia uniparental.

Esse fenômeno quebra a regra de que as crianças afetadas por um distúrbio recessivo aparecem apenas em famílias em que ambos os genitores são portadores. Por exemplo, a fibrose cística é uma doença recessiva autossômica; tipicamente, o pai e a mãe de uma criança afetada são heterozigotos para a mutação da fibrose cística no cromossomo 7. Entretanto, para uma pequena proporção de pessoas com fibrose cística, apenas um dos pais é heterozigoto para essa mutação. Nesses casos, a fibrose cística surge graças à dissomia uniparental: a pessoa que tem fibrose cística herdou do genitor heterozigoto duas cópias do cromossomo 7 com o alelo defeituoso dessa doença do cromossomo 7 e nenhuma cópia do alelo normal do outro genitor.

A dissomia uniparental também foi observada na síndrome de Prader-Willi, uma condição rara que surge quando não existe uma cópia paterna de um gene no cromossomo 15. Embora a maioria dos casos da síndrome de Prader-Willi surja a partir de uma deleção cromossômica que remove a cópia paterna do gene (ver *Capítulo 5*), 20 a 30% dos casos surgem quando ambas as cópias do cromossomo 15 são herdadas da mãe e nenhuma cópia é herdada do pai.

Mosaicismo

A não disjunção em uma divisão mitótica pode gerar regiões de células nas quais todas as células têm uma anormalidade cromossômica, e outras regiões nas quais todas as células têm um cariótipo normal. Esse tipo de não disjunção forma regiões do tecido com diferentes composições cromossômicas, uma condição conhecida como **mosaicismo**. Evidências crescentes sugerem que o mosaicismo é mais comum do que é identificado. Por exemplo, cerca de 50% das pessoas diagnosticadas com síndrome de Turner (indivíduos com um único cromossomo X) são, na verdade, mosaicos, com alguns casos com células 45,X e alguns normais com células 46,XX. Alguns são até mosaicos para dois ou mais tipos de cariótipos anormais. O mosaico 45,X/46,XX em geral surge quando um cromossomo X é perdido logo após a fertilização em um embrião XX.

As moscas-da-fruta que são mosaicos XX/XO (O indica a ausência de um cromossomo homólogo; XO significa que a célula tem um único cromossomo X e nenhum cromossomo Y) desenvolvem uma mistura de traços de machos e fêmeas, porque a existência de dois cromossomos X nessas moscas produz traços de fêmeas e a de um único cromossomo X produz traços de machos (**Figura 8.23**). Nas moscas-da-fruta, o sexo é determinado independentemente em cada célula no curso do desenvolvimento. As células que são XX expressam traços de fêmeas; as que são XO expressam traços de machos. Esses mosaicos sexuais são chamados de **ginandromorfos**. Normalmente, os genes recessivos ligados ao X são mascarados nas fêmeas heterozigotas, mas, nos mosaicos XX/XO, qualquer gene recessivo ligado ao X nas células com um único cromossomo X será expresso.

Conceitos

Na dissomia uniparental, um organismo tem duas cópias de um cromossomo de um genitor e nenhuma cópia do outro. Essa dissomia surge quando um embrião trissômico perde um dos cromossomos triplicados no início do desenvolvimento. No mosaicismo, diferentes células do mesmo organismo têm diferentes composições cromossômicas.

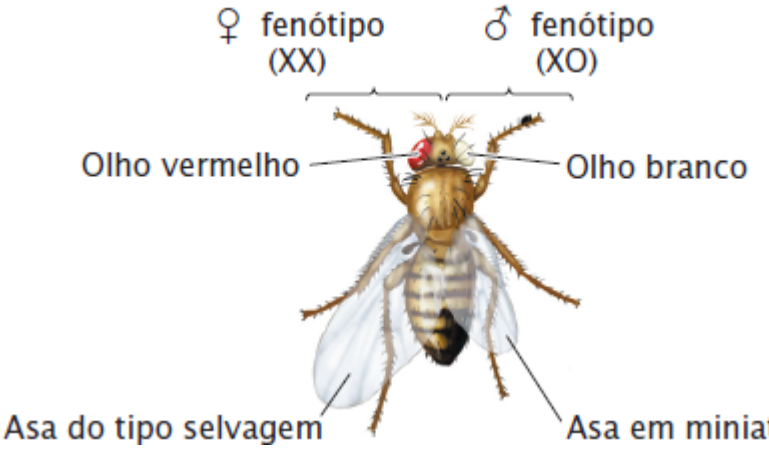


Figura 8.23 O mosaicismo para os cromossomos sexuais produz um ginandromorfo. Essa mosca-da-fruta ginandromorfa XX/XO carreia um cromossomo X do tipo selvagem e um cromossomo X com alelos recessivos para olhos brancos e asas em miniatura. O lado esquerdo da mosca tem um fenótipo normal de fêmea, porque as células são XX e os alelos recessivos no cromossomo X são mascarados pela existência dos alelos do tipo selvagem no outro. O lado direito da mosca tem um fenótipo de macho com olhos brancos e asa em miniatura, porque as células não têm o cromossomo X do tipo selvagem (são XO), permitindo que os alelos branco e miniatura sejam expressos.

8.4 Poliploidia é a ocorrência de mais de dois conjuntos de cromossomos

Como já foi discutido na introdução deste capítulo, alguns organismos (como as bananas) têm mais de dois conjuntos de cromossomos e são poliploides. Os poliploides incluem *triploides* (3*n*), *tetraploides* (4*n*), *pentaploides* (5*n*) e até números maiores de conjuntos de cromossomos.

A poliploidia é comum nas plantas e é o principal mecanismo pelo qual surgem novas espécies de plantas. Aproximadamente 40% de todas as espécies de plantas com fluorescência e 70 a 80% da grama são poliploides. Elas incluem várias plantas importantes para a agricultura, como trigo, aveia, algodão, batatas e cana-de-açúcar. A poliploidia é menos comum nos animais, mas é encontrada em alguns invertebrados, peixes, salamandras, sapos e lagartos. Embora não ocorram naturalmente, os poliploides viáveis são conhecidos nas aves, mas foi relatado pelo menos um mamífero poliploide – um rato na Argentina.

Vamos considerar os dois principais tipos de poliploidia: **autopoliploidia**, na qual todos os conjuntos de cromossomos são de uma única espécie, e **alopoliploidia**, na qual os conjuntos de cromossomos são oriundos de duas ou mais espécies.

Autopoliploidia

A autopoliploidia é provocada por acidentes da mitose ou meiose que produzem conjuntos extras de cromossomos, todos derivados de uma única espécie. A não disjunção de todos os cromossomos na mitose em um embrião 2*n* inicial, por exemplo, dobra o número de cromossomos e produz um autotetraploide (4*n*), como descrito na **Figura 8.24 A**. Um autotriploide (3*n*) pode surgir quando a não disjunção na meiose produz um gameta diploide que então se une a um gameta haploide normal para produzir um zigoto triploide (**Figura 8.24 B**). Por outro lado, os triploides podem surgir a partir de um cruzamento entre um autotetraploide que produz gametas 2*n* e um diploide que produz gametas 1*n*. A não disjunção pode ser artificialmente induzida pela colchicina, uma substância química que interrompe a formação do fuso. A colchicina é, com frequência, usada para induzir poliploidia nas plantas importantes para agricultura e decoração.

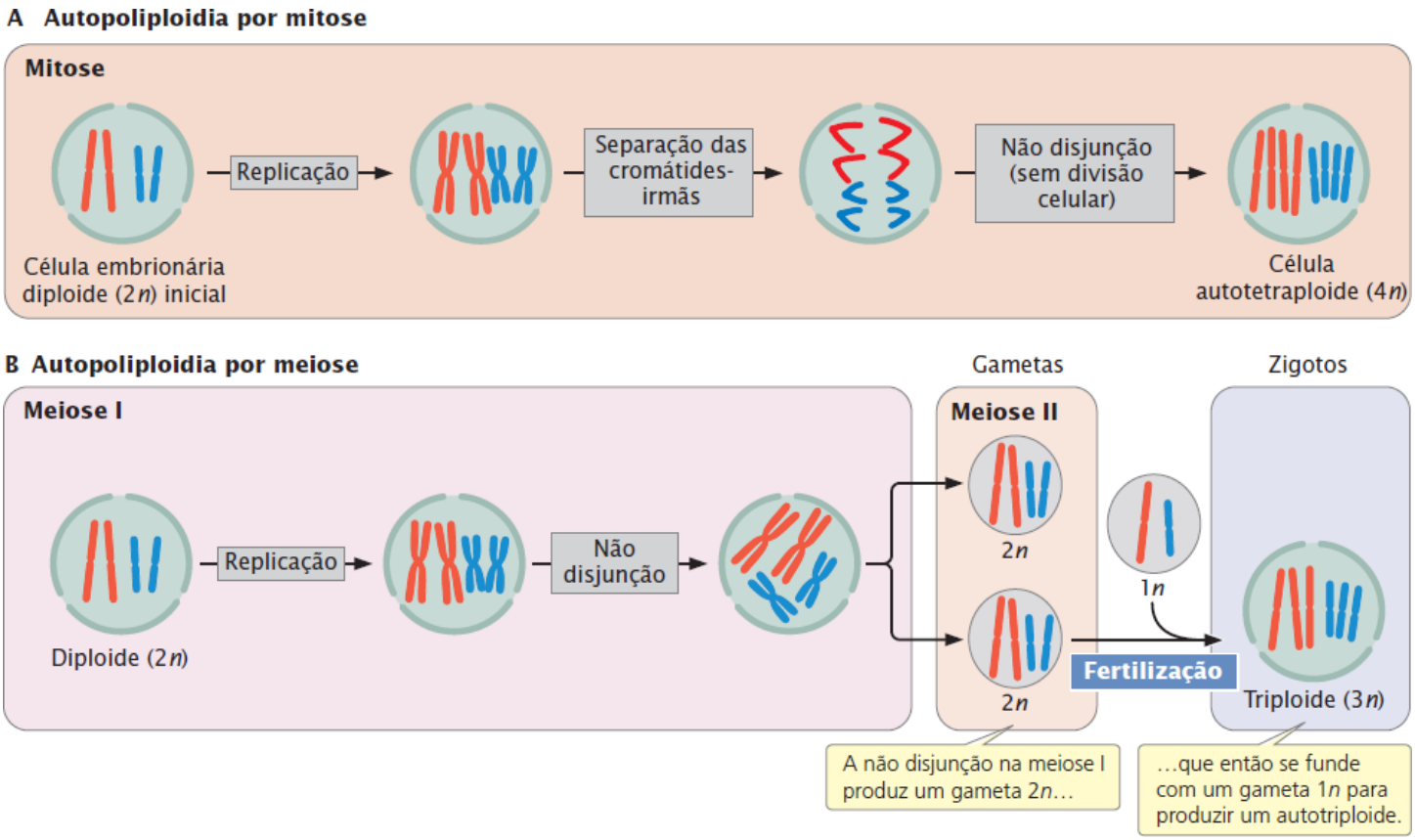


Figura 8.24 **A autopoliploidia pode surgir pela não disjunção na mitose ou meiose.**

Como todos os conjuntos de cromossomos nos autopoliploides são da mesma espécie, eles são homólogos e tentam se alinhar na prófase I da meiose, que em geral resulta em esterilidade. Considere a meiose em um autotriploide (**Figura 8.25**). Na meiose em uma célula diploide, dois cromossomos homólogos se pareiam e se alinham, mas, nos autotriploides, três homólogos estão presentes. Um dos três homólogos pode não conseguir se alinhar com os outros dois, e esse cromossomo não alinhado se separa aleatoriamente (ver Figura 8.25 A). O gameta que conseguirá o cromossomo extra será determinado por acaso e será diferente para cada grupo homólogo de cromossomos. Os gametas resultantes terão duas cópias de alguns cromossomos e uma cópia de outros. Mesmo se todos os três cromossomos se alinharem, dois cromossomos têm de se separar para um gameta e um cromossomo para o outro (ver Figura 8.25 B). Às vezes, a existência de um terceiro cromossomo interfere no alinhamento normal, e todos os três se deslocam para o mesmo gameta (ver Figura 8.25 C).

Não importa como os três cromossomos homólogos se alinham, sua separação aleatória criará **gametas não balanceados**, com vários números de cromossomos. Um gameta produzido por meiose nesse autotriploide pode receber, por assim dizer, duas cópias do cromossomo 1, uma cópia do cromossomo 2, três cópias do cromossomo 3 e nenhuma cópia do cromossomo 4. Quando o gameta não balanceado se une a um gameta normal (ou com outro gameta não equilibrado) o zigoto resultante tem números diferentes dos quatro tipos de cromossomos. Essa diferença no número cria uma dosagem gênica não balanceada no zigoto, que é fatal. Por esse motivo, os triploides não produzem descendentes viáveis.

Nos autopoliploides de número par, como os autotetraploides, os cromossomos homólogos podem teoricamente formar pares e se dividir igualmente. Entretanto, é raro esse evento acontecer, então esses tipos de autotetraploides também produzem gametas não balanceados.

A esterilidade que, em geral, acompanha a autopoliploidia é explorada na agricultura. Como foi discutido na introdução deste capítulo, as bananas triploides ($3n = 33$) são estéreis e sem semente. Da mesma forma, foram criadas melancias triploides sem sementes, que são amplamente vendidas.

Alopoliploidia

A alopoliploidia surge a partir da hibridização entre duas espécies; o poliploide resultante carreia conjuntos de cromossomos derivados de duas ou mais espécies. A **Figura 8.26** mostra como a alopoliploidia pode surgir a partir de duas espécies que são próximas o suficiente para que ocorra hibridização entre elas. A espécie 1 (AABBCC, $2n = 6$) produz gametas haploides com os cromossomos ABC, e a espécie 2 (GGHHII, $2n = 6$) produz gametas com os cromossomos GHI. Se os gametas das espécies 1 e 2 se unem, é criado um híbrido com seis cromossomos (ABCGHI). O híbrido tem o mesmo número de cromossomos que ambas as espécies diploides, então é considerado um diploide. Entretanto, como os cromossomos híbridos não são homólogos, eles não pareiam e se separam corretamente na meiose; esse híbrido é haploide e estéril do ponto de vista funcional.

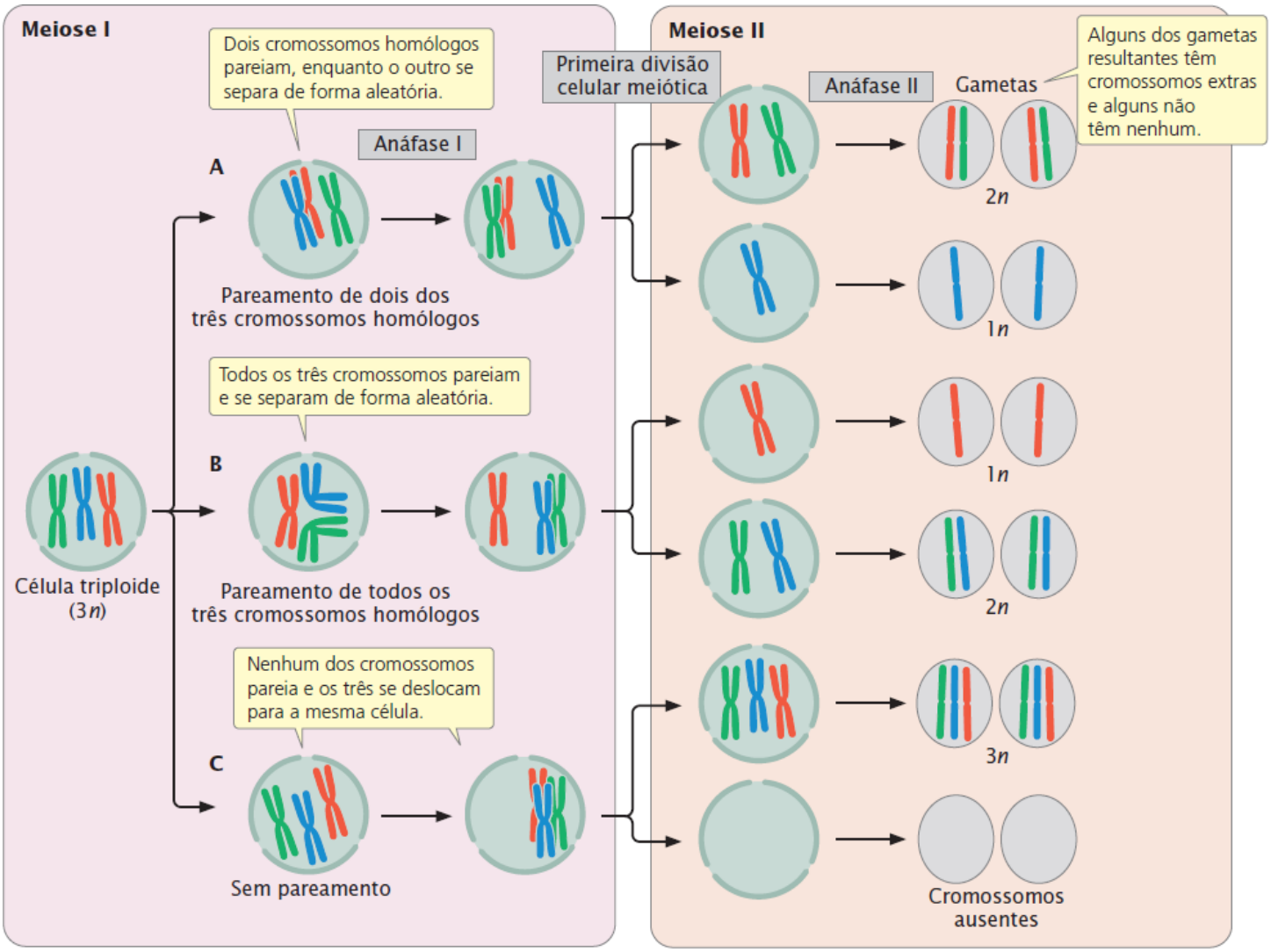
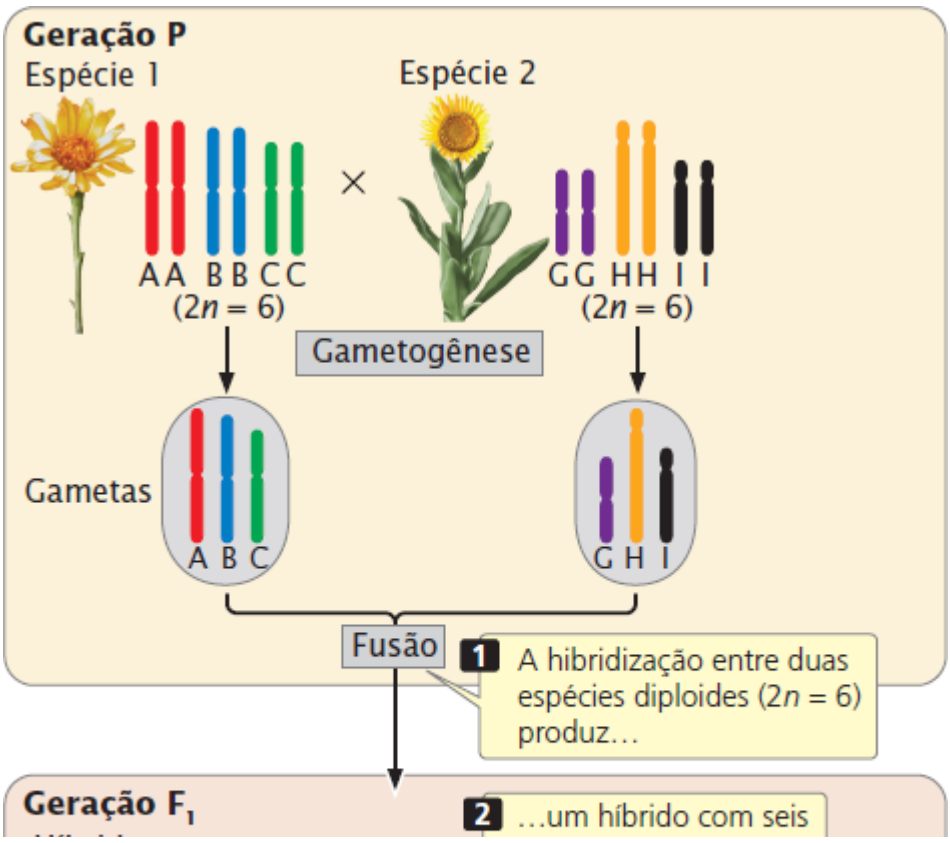


Figura 8.25 Na meiose de um autotriploide, os cromossomos homólogos podem parear ou não de três formas. Este exemplo ilustra o pareamento e a separação de um único conjunto homólogo de cromossomos.



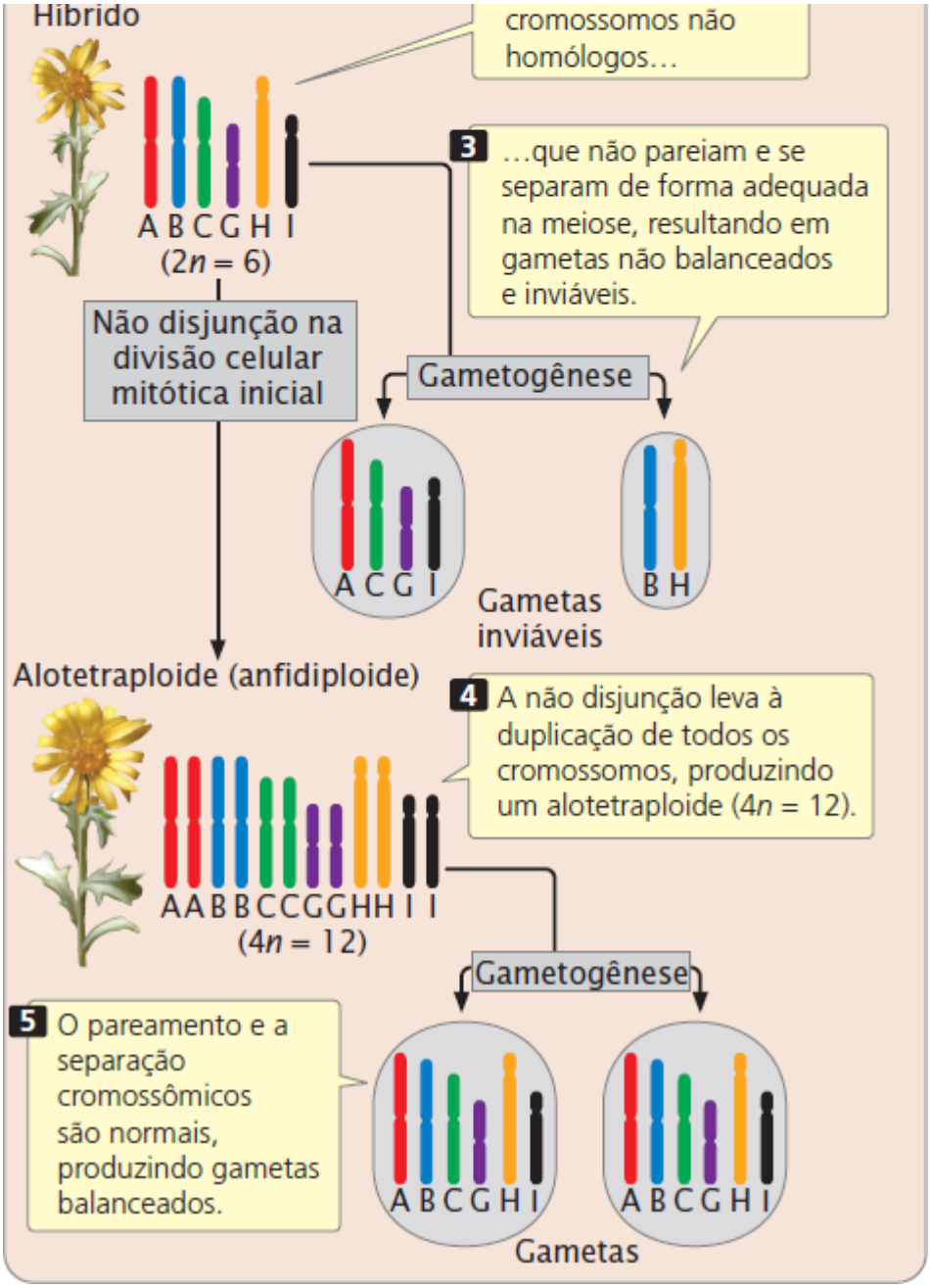


Figura 8.26 A maioria dos alopoliploides surge a partir da hibridização entre duas espécies seguida por duplicação do cromossomo.

O híbrido estéril não produz gametas viáveis por meiose, mas pode ser capaz de se perpetuar por mitose (reprodução assexuada). Em raras ocasiões, a não disjunção ocorre em uma divisão mitótica, que leva a dobrar o número de cromossomos e a um alotetraploide com cromossomos AABBCCGGHHII. Esse tipo de alopoliploide, com dois genomas diploides combinados, é às vezes chamado de **anfidiploide**. Embora o número de cromossomos tenha dobrado em comparação com o existente em cada uma das espécies genitoras, o anfidiploide é funcionalmente diploide: cada cromossomo tem um, e apenas um, parceiro homólogo, que é exatamente o que a meiose requer para a separação adequada. O anfidiploide agora pode sofrer meiose normal para produzir gametas balanceados com seis cromossomos.

George Karpechenko criou poliploides experimentalmente nos anos 1920. O repolho (*Brassica oleracea*, $2n = 18$) e o rabanete (*Raphanus sativa*, $2n = 18$) são plantas importantes para a agricultura, mas apenas as folhas do repolho e as raízes do rabanete são consumidas. Karpechenko queria produzir uma planta que tivesse as folhas do repolho e as raízes do rabanete, assim nenhuma parte da planta seria desperdiçada. Como o repolho e o rabanete têm 18 cromossomos, Karpechenko conseguiu cruzá-los com sucesso, produzindo um híbrido com $2n = 18$, mas, infelizmente, o híbrido era estéril. Após vários cruzamentos, Karpechenko observou que uma de suas plantas produziu algumas sementes. Quando plantadas, elas cresceram em plantas viáveis e férteis. A análise dos seus cromossomos revelou que as plantas eram alotetraploides, com $2n = 36$ cromossomos. Para decepção de Karpechenko, entretanto, as novas plantas tinham as raízes do repolho e as folhas do rabanete.

Problema desenvolvido

A espécie I tem 2*n* = 14 e a espécie II tem 2*n* = 20. Indique todos os possíveis números de cromossomos encontrados nos indivíduos a seguir.

a. Um autotriploide da espécie I.
b. Um autotetraploide da espécie II.
c. Um alotriploide formado a partir das espécies I e II.
d. Um alotetraploide formado a partir das espécies I e II.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

Todos os possíveis números de cromossomos para os indivíduos com o tipo de poliploidia indicado.

Quais informações são fornecidas para solucionar o problema?

- A espécie I tem 2*n* = 14 e a espécie II tem 2*n* = 20
- O tipo de poliploidia que o indivíduo tem.

Etapas para a solução

O número de cromossomos haploides (*n*) para a espécie I é 7, e para espécie II é 10.

a. Um indivíduo triploide é 3*n*. Um erro comum é considerar que 3*n* significa três vezes o número de cromossomos de um indivíduo normal, mas lembre-se de que os indivíduos normais são 2*n*. Uma vez que *n* para a espécie I é 7 e todos os genomas de um autopoliploide são da mesma espécie, 3*n* = 3 × 7 = 21.
b. Um autotetraploide é 4*n* com todos os genomas da mesma espécie. O *n* para a espécie II é 10, então 4*n* = 4 × 10 = 40.
c. Um triploide é 3*n*. Por definição, um alopoliploide tem obrigatoriamente genomas de duas espécies diferentes. Um alotriploide poderia ter 1*n* da espécie I e 2*n* da espécie II ou (1 × 7) + (2 × 10) = 27. Por outro lado, ele poderia ter 2*n* da espécie I e 1*n* da espécie II ou (2 × 7) + (1 × 10) = 24. Assim, o número de cromossomos em um alotriploide pode ser 24 ou 27.
d. Um tetraploide é 4*n*. Por definição, um alotetraploide tem obrigatoriamente genomas de pelo menos duas espécies diferentes. Um alotetraploide poderia ter 3*n* a partir da espécie I e 1*n* da espécie II ou (3 × 7) + (1 × 10) = 31; ou 2*n* da espécie I e 2*n* da espécie II ou (2 × 7) + (2 × 10) = 34; ou 1*n* da espécie I e 3*n* da espécie II ou (1 × 7) + (3 × 10) = 37. Assim, o número de cromossomos poderia ser 31, 34 ou 37.

Resolva o **Problema 38** no final deste capítulo para exercitar mais.

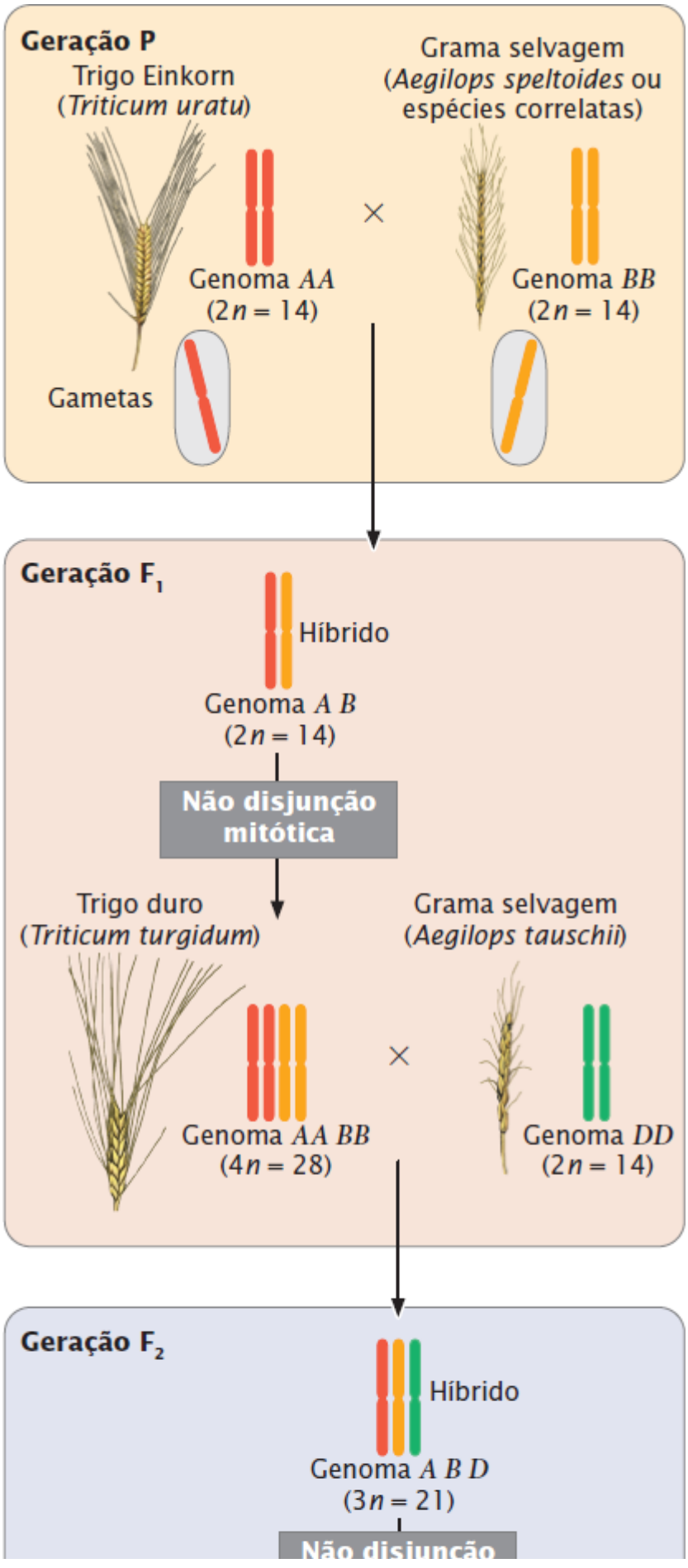
Importância da poliploidia

Em muitos organismos, o volume celular está relacionado com o volume do núcleo, que, por sua vez, é determinado pelo tamanho do genoma. Assim, o aumento no número de cromossomos na poliploidia está associado a aumento do tamanho da célula, e muitos poliploides são maiores que os diploides. Os produtores usaram esse efeito para produzir plantas com folhas, flores, frutas e sementes maiores. É provável que o genoma do hexaploide (6*n* = 42) do trigo tenha cromossomos derivados de três espécies selvagens diferentes (**Figura 8.27**). Assim, as sementes do trigo moderno são maiores que as sementes de seus ancestrais. Muitas outras plantas cultivadas também são poliploides (**Quadro 8.2**).

A poliploidia é menos comum em animais do que em plantas por vários motivos. Como já foi discutido, os alopoliploides exigem a hibridização entre diferentes espécies, o que é menos frequente em animais do que em plantas. O comportamento animal evita o intercruzamento entre espécies, e a complexidade do desenvolvimento animal torna inviáveis híbridos mais interespecíficos. Muitos dos animais poliploides que surgiram estão em grupos que se reproduzem pela partenogênese (um tipo de reprodução na qual o animal se desenvolve a partir de um óvulo não fertilizado). Assim, a reprodução assexuada facilita o desenvolvimento de poliploides, talvez porque a perpetuação de híbridos por reprodução assexuada forneça maiores oportunidades para a não disjunção do que a reprodução sexuada. Foram relatados apenas alguns recém-nascidos humanos poliploides, e a maioria morreu com alguns dias de vida. Poliploidia – em geral triploidia – é observada em cerca de 10% dos fetos humanos abortados espontaneamente.

Importância da poliploidia na evolução. A poliploidia, em especial a alopoliploidia, dá origem a novas espécies e é particularmente importante na evolução das plantas florescentes. A duplicação genômica por poliploidia é o principal contribuinte para o sucesso da evolução nos vários grupos. Por exemplo, *Saccharomyces cerevisiae* (levedura) é um tetraploide

que sofreu duplicação do genoma inteiro há cerca de 100 milhões de anos. O genoma dos vertebrados duplicou duas vezes, uma no ancestral comum dos vertebrados com mandíbulas e novamente no ancestral dos peixes. Alguns grupos de vertebrados, como determinados sapos e peixes, sofreram poliploidia adicional. Os cereais sofreram vários eventos de duplicação do genoma. Os diferentes tipos de mutações cromossômicas estão resumidos no **Quadro 8.3**.



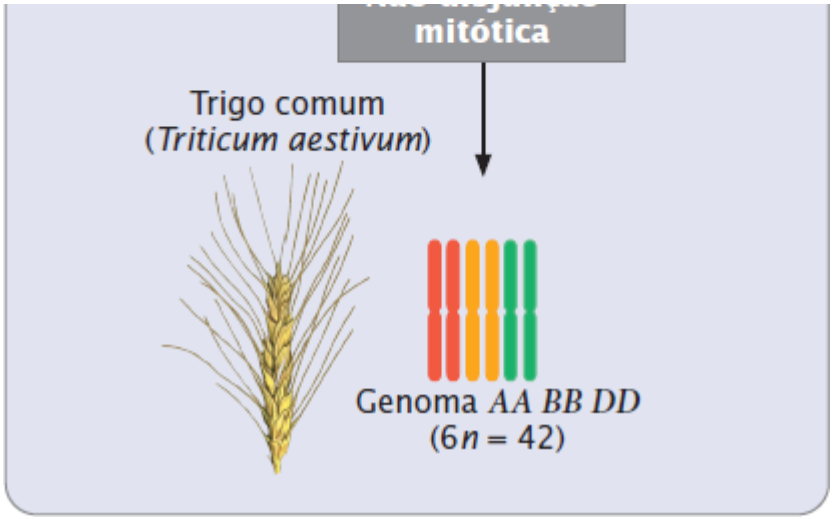


Figura 8.27 O trigo comum moderno, *Triticum aestivum*, é um hexaploide com genes derivados de três espécies diferentes. Duas espécies diploides, *T. uratu* (n = 14) e provavelmente *Aegilops speltoides* ou uma espécie relacionada (n = 14), originalmente cruzadas para produzir um híbrido diploide ($2n$ = 14) que sofreu duplicação cromossômica para criar *T. turgidum* ($4n$ = 28). Um cruzamento entre *T. turgidum* e *A. tauschii* ($2n$ = 14) produziu um híbrido diploide ($3n$ = 21) que então sofreu duplicação cromossômica para produz *T. aestivum*, que é hexaploide ($6n$ = 42).

Quadro 8.2 Exemplos de plantas poliploides cultivadas.

Planta	Tipo de poliploidia	Ploidia	Número de cromossomos
Batata	Autopoliploide	4*n*	48
Banana	Autopoliploide	3*n*	33
Amendoim	Autopoliploide	4*n*	40
Batata-doce	Autopoliploide	6*n*	90
Tabaco	Alopoliploide	4*n*	48
Algodão	Alopoliploide	4*n*	52
Trigo	Alopoliploide	6*n*	42
Aveia	Alopoliploide	6*n*	42
Cana-de-açúcar	Alopoliploide	8*n*	80
Morango	Alopoliploide	8*n*	56

Fonte: De F. C. Elliot, Plant Breeding and Cytogenetics (New York: McGraw-Hill, 1958).

Conceitos

A poliploidia é a existência de conjuntos extras de cromossomos: os autopoliploides têm conjuntos extras de cromossomos a partir da mesma espécie; os alopoliploides têm conjuntos extras de cromossomos a partir de duas ou mais espécies. Os problemas de pareamento e separação cromossômicos levam à esterilidade em autopoliploides, mas muitos alopoliploides são férteis.

Checagem dos conceitos 7

A espécie A tem $2n$ = 16 cromossomos, e a espécie B tem $2n$ = 14. Quantos cromossomos seriam encontrados em um alotriploide dessas duas espécies?

a. 21 ou 24.
b. 42 ou 48.

c. 22 ou 23.
d. 45.

Quadro 8.3 Diferentes tipos de mutações cromossômicas.

Mutação cromossômica	Definição
Rearranjo cromossômico	Alteração na estrutura do cromossomo
Duplicação cromossômica	Duplicação de um segmento de cromossomo
Deleção cromossômica	Deleção de um segmento de cromossomo
Inversão	Segmento de cromossomo invertido em 180º
Inversão paracêntrica	Inversão que não inclui o centrômero na região invertida
Inversão pericêntrica	Inversão que inclui o centrômero na região invertida
Translocação	Movimento de um segmento do cromossomo para um cromossomo não homólogo ou para outra região do mesmo cromossomo
Translocação não recíproca	Movimento de um segmento do cromossomo para um cromossomo não homólogo ou para outra região do mesmo cromossomo sem troca recíproca
Translocação recíproca	Troca entre segmentos de cromossomos não homólogos ou entre regiões do mesmo cromossomo
Aneuploidia	Mudança no número de cromossomos individuais
Nulissomia	Perda de ambos os membros de um par homólogo
Monossomia	Perda de um membro de um par homólogo
Trissomia	Ganho de um cromossomo, resultando em três cromossomos homólogos
Tetrassomia	Ganho de dois cromossomos homólogos, resultando em quatro cromossomos homólogos
Poliploidia	Adição de conjuntos inteiros de cromossomos
Autopoliploidia	Poliploidia na qual conjuntos extras de cromossomos provêm da mesma espécie
Alopoliploidia	Poliploidia na qual conjuntos extras de cromossomos provêm de duas ou mais espécies

Resumo dos conceitos

- Os três tipos básicos de mutações cromossômicas são: (1) rearranjos cromossômicos, que são mudanças nas estruturas dos cromossomos; (2) aneuploidia, que é o aumento ou a redução no número de cromossomos e (3) poliploidia, a existência de conjuntos extras de cromossomos
- Os rearranjos cromossômicos incluem duplicações, deleções, inversões e translocações
- Nos indivíduos heterozigotos para uma duplicação, uma região duplicada formará uma alça quando os cromossomos homólogos pareiam na meiose. As duplicações têm efeitos significativos no fenótipo, possivelmente ao desequilibrar a dosagem gênica. As duplicações segmentares são comuns no genoma humano
- Nos indivíduos heterozigotos para uma deleção, um dos cromossomos forma uma alça durante o pareamento na meiose. As deleções podem fazer com que os alelos recessivos sejam expressos
- As inversões pericêntricas incluem o centrômero; as inversões paracêntricas, não. Nos indivíduos heterozigotos para uma inversão, os cromossomos homólogos formam alças de inversão na meiose, com a recombinação reduzida ocorrendo dentro

da região invertida

- Nos heterozigotos para translocação, os cromossomos formam estruturas semelhantes à cruz na meiose, e a separação dos cromossomos produz gametas não balanceados
- Os sítios frágeis são constrições ou intervalos que surgem em regiões particulares nos cromossomos das células que crescem em cultura e são suscetíveis a ruptura sob algumas condições
- As variações no número de cópias (CNVs) são diferenças no número de cópias de sequências de DNA e incluem as duplicações e as deleções. Essas variantes são comuns no genoma humano; algumas estão associadas a doenças e distúrbios
- A nulissomia é a perda de dois cromossomos homólogos; a monossomia é a perda de um cromossomo homólogo; a trissomia é a adição de um cromossomo homólogo; a tetrassomia é a adição de dois cromossomos homólogos
- A aneuploidia, em geral, provoca efeitos fenotípicos drásticos porque desequilibra a dosagem gênica
- A síndrome de Down primária é causada pela existência de três cópias completas do cromossomo 21, enquanto a síndrome de Down familiar é causada por duas cópias normais do cromossomo 21 e uma terceira cópia que está ligada a outro cromossomo por meio de uma translocação
- A dissomia uniparental é a existência de duas cópias de um cromossomo de um genitor e nenhuma cópia do outro. O mosaicismo é causado pela não disjunção em uma divisão mitótica inicial que leva a diferentes composições cromossômicas em diferentes células em um único indivíduo
- Todos os cromossomos em um autopoliploide provêm de uma espécie; os cromossomos em um alopoliploide provêm de duas ou mais espécies.

Termos importantes

Alopoliploidia
Aneuploidia
Anfidiploidia
Autopoliploidia
Cromátide acêntrica
Cromátide dicêntrica
Cromossomo acrocêntrico
Cromossomo metacêntrico
Cromossomo submetacêntrico
Cromossomo telocêntrico
Deleção cromossômica
Dissomia uniparental
Duplicação cromossômica
Duplicação deslocada
Duplicação em *tandem*
Duplicação reversa
Duplicação segmentar
Efeito da posição
Gametas não balanceados
Gene haploinsuficiente
Ginandromorfo
Inversão cromossômica
Inversão paracêntrica
Inversão pericêntrica
Monossomia
Mosaicismo
Mutação cromossômica
Nulissomia
Poliploidia
Ponte dicêntrica
Portador da translocação
Pseudodominância
Rearranjo cromossômico
Segregação adjacente-1

Segregação adjacente-2
Segregação alternada
Síndrome de Down (trissomia do 21)
Síndrome de Down familiar
Síndrome de Down primária
Síndrome de Edward (trissomia do 18)
Síndrome de Patau (trissomia do 13)
Síndrome do X frágil
Sítio frágil
Tetrassomia
Translocação
Translocação não recíproca
Translocação recíproca
Translocação robertsoniana
Trissomia
Trissomia do 8
Variação no número de cópias (CNV)

Respostas da Checagem dos conceitos

1. a.
2. A pseudodominância é a expressão de uma mutação recessiva. É produzida quando o alelo dominante do tipo selvagem em um indivíduo heterozigoto não existe por causa de uma deleção em um cromossomo.
3. c.
4. b.
5. 37.
6. A compensação da dosagem evita a expressão de cópias adicionais de genes ligados ao X nos mamíferos e existe pouca informação no cromossomo Y, então cópias extras dos cromossomos X e Y não têm efeitos significativos no desenvolvimento. Por outro lado, não existe mecanismo de compensação de dosagem para os autossomos, e as cópias extras de genes autossômicos são expressas, atrapalhando o desenvolvimento e provocando o aborto espontâneo de embriões aneuploides.
7. c.

Problemas desenvolvidos

Problema 1

Um cromossomo tem os seguintes segmentos, em que • representa o centrômero.

A B C D E • F G

Que tipos de mutações cromossômicas são necessários para mudar esse cromossomo em cada um dos cromossomos a seguir? (Em alguns casos, é necessária mais de uma mutação cromossômica.)

a. A B E • F G
b. A E D C B • F G
c. A B A B C D E • F G
d. A F • E D C B G
e. A B C D E E D C • F G

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Os tipos de mutações cromossômicas que levariam ao cromossomo apresentado.

Quais informações são fornecidas para solucionar o problema?

- Os segmentos de genes originais encontrados no cromossomo

- Os segmentos de gene alterados que ocorrem após as mutações.

Para a solução deste problema, revise:
Seção 8.2.

Etapas para a solução

a. O cromossomo mutante (A B E • F G) não tem o segmento C D; então essa mutação é uma deleção.
b. O cromossomo mutante (A E D C B • F G) tem apenas uma cópia de todos os segmentos, mas o segmento B D C E está invertido em 180º. Como a localização do centrômero não foi alterada e não está na região invertida, essa mutação é uma inversão paracêntrica.
c. O cromossomo mutante (A B A B C D E • F G) é maior que o normal, e vemos que o segmento A B foi duplicado. Essa mutação é uma duplicação em *tandem*.
d. O cromossomo mutante (A F • E D C B G) tem o tamanho normal, mas a ordem dos genes e a localização do centrômero foram alteradas; essa mutação é uma inversão pericêntrica da região (B C D E • F).
e. O cromossomo mutante (A B C D E E D C • F G) tem uma duplicação (C D E) que também está invertida, então este cromossomo sofreu uma duplicação e uma inversão paracêntrica.

Problema 2

A espécie I é diploide ($2n = 4$) com os cromossomos AABB; a espécie II é diploide ($2n = 6$) com os cromossomos MMNNOO. Indique os cromossomos que seriam encontrados nos indivíduos com as seguintes mutações cromossômicas.

a. Autotriploidia na espécie I.
b. Alotetraploidia incluindo as espécies I e II.
c. Monossomia na espécie I.
d. Trissomia na espécie II para o cromossomo M.
e. Tetrassomia na espécie I para o cromossomo A.
f. Alotriploidia incluindo as espécies I e II.
g. Nulissomia na espécie I para o cromossomo N.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
As designações de letras dos cromossomos que serão encontrados nos indivíduos com cada tipo de mutação.

Quais informações são fornecidas para solucionar o problema?

- A espécie I é diploide com $2n = 4$
- A espécie I tem os cromossomos AABB
- A espécie II é diploide com $2n = 6$
- A espécie II tem MMNNOO.

Para a solução deste problema, revise:
Seções 8.3 e 8.4.

Etapas para a solução

a. Um autotriploide é $3n$, com todos os cromossomos surgindo a partir de uma única espécie; então, um autotriploide da espécie I teria os cromossomos AAABBB ($3n = 6$).

Dica: Determine primeiro o complemento do genoma haploide para cada espécie. Para a espécie I, $n = 2$ com os cromossomos AB e, para a espécie II, $n = 3$ com os cromossomos MNO.

b. Um alotetraploide é 4 n, com os cromossomos vindo de mais de uma espécie. Um alotetraploide pode ter $2n$ da espécie I e $2n$ da espécie II, gerando o alotetraploide ($4n = 2 + 2 + 3 + 3 = 10$) com os cromossomos AABBMMNNOO. Um

alotetraploide também pode ter 3*n* da espécie I e 1 *n* da espécie II (4 *n* = 2 + 2 + 2 + 3 = 9; AAABBBMNO) ou 1*n* da espécie I e 3*n* da espécie II (4 *n* = 2 + 3 + 3 + 3 = 11, ABMMMNNNOOO).

c. Um monossômico tem apenas um cromossomo ausente; então, um monossômico para a espécie I seria 2*n* − 1 = 4 − 1 = 3. A monossomia pode incluir um dos dois pares de cromossomos, com os cromossomos ABB ou AAB.
d. A trissomia requer um cromossomo extra; desse modo, um trissômico da espécie II para o cromossomo M seria 2*n* + 1 = 6 + 1 = 7 (MMMNNOO).
e. Um tetrassômico tem dois cromossomos homólogos extras; então, um tetrassômico da espécie I para o cromossomo A seria 2*n* + 2 = 4 + 2 = 6 (AAAABB).
f. Um alotriploide é 3*n* com os cromossomos oriundos de duas espécies diferentes; assim, um alotriploide seria 3*n* = 2 + 2 + 3 = 7 (AABBMNO) ou 3*n* = 2 + 3 + 3 = 8 (ABMMNNOO).
g. Um nulissômico não tem ambos os cromossomos de um par homólogo; assim, um nulissômico da espécie II para o cromossomo N seria 2*n* − 2 = 6 − 2 = 4 (MMOO).

Questões de compreensão

Seção 8.1

1. Liste os diferentes tipos de mutações cromossômicas e defina cada uma delas.

Seção 8.2

2. Por que as cópias extras de genes provocam efeitos fenotípicos drásticos?
3. Desenhe um par de cromossomos como ele apareceria durante a sinapse na prófase I da meiose de um indivíduo heterozigoto para uma duplicação cromossômica.
4. O que é haploinsuficiência?
5. Qual é a diferença entre uma inversão paracêntrica e uma inversão pericêntrica?
6. Como as inversões nas quais não há ganho nem perda de informação genética podem causar efeitos fenotípicos?
7. Explique, com a ajuda de um desenho, como uma ponte dicêntrica é produzida quando o *crossing over* ocorre em um heterozigoto para uma inversão paracêntrica.
8. Explique por que a recombinação é suprimida em indivíduos heterozigotos para as inversões paracêntricas e pericêntricas.
9. Como as translocações nas quais não há ganho nem perda de informação genética causam efeitos fenotípicos?
10. Esboce o pareamento cromossômico e os diferentes padrões de separação que podem surgir em um indivíduo heterozigoto para uma translocação recíproca.
11. O que é uma translocação robertsoniana?

Seção 8.3

12. Liste os quatro principais tipos de aneuploidia.
13. Qual é a diferença entre a síndrome de Down primária e a síndrome de Down familiar? Como surge cada um desses tipos?
14. O que é dissomia uniparental e como ela surge?
15. O que é mosaicismo e como ele surge?

Seção 8.4

16. Qual é a diferença entre autopoliploidia e alopoliploidia? Como surge cada uma?
17. Explique por que os autopoliploides, em geral, são estéreis, enquanto os alopoliploides são férteis.

Seção 8.1

18. Examine os cariótipos apresentados nas Figuras 8.1 e 8.2. Os indivíduos a partir dos quais foram feitos esses cariótipos eram machos ou fêmeas?

*19. Que tipos de mutações cromossômicas:
    a. aumentam a quantidade de material genético em um determinado cromossomo?
    b. aumentam a quantidade de material genético em todos os cromossomos?
    c. reduzem a quantidade de material genético em determinado cromossomo?
    d. alteram a posição das sequências de DNA em um único cromossomo sem alterar a quantidade de material genético?

e. deslocam o DNA de um cromossomo para um cromossomo não homólogo?

Seção 8.2

*20. Um cromossomo tem os segmentos a seguir, em que • representa o centrômero:

A B • C D E F G

Que tipos de mutações cromossômicas são necessários para mudar esse cromossomo em cada um dos cromossomos a seguir? (Em alguns casos, é necessária mais de uma mutação cromossômica.)

a. A B A B • C D E F G
b. A B • C D E A B F G
c. A B • C F E D G
d. A • C D E F G
e. A B • C D E
f. A B • E D C F G
g. C • B A D E F G
h. A B • C F E D F E D G
i. A B • C D E F C D F E G

21. Um cromossomo inicialmente tem os segmentos a seguir:

A B • C D E F G

Desenhe o cromossomo, identificando seus segmentos, que seriam resultado para cada uma das mutações a seguir.

a. Duplicação em *tandem* de DEF.
b. Duplicação deslocada de DEF.
c. Deleção de FG.
d. Inversão paracêntrica que inclui DEFG.
e. Inversão pericêntrica de BCDE.

22. O diagrama a seguir representa dois cromossomos não homólogos:

A B • C D E F G
R S • T U V W X

Que tipo de mutação cromossômica produziria cada um dos cromossomos a seguir?

a. A B • C D R S • T U V W X E F G
b. A U V B • C D E F G R S • T W X
c. A B • T U V F G R S • C D E W X
d. A B • C W G R S • T U V D E F X

*23. A mutação *Notch* é uma deleção no cromossomo X da *Drosophila melanogaster*. As moscas-fêmeas heterozigotas para *Notch* têm um entalhe nas margens de suas asas; a mutação *Notch* é fatal nas condições homozigotas e hemizigotas. A deleção *Notch* cobre a região do cromossomo X que tem o *locus* para olhos brancos, um traço recessivo ligado ao X. Indique os fenótipos e as proporções dos descendentes produzidos pelos seguintes cruzamentos.

a. Uma fêmea *Notch* de olhos vermelhos é cruzada com um macho de olhos brancos.
b. Uma fêmea *Notch* de olhos brancos é cruzada com um macho de olhos vermelhos.
c. Uma fêmea *Notch* de olhos brancos é cruzada com um macho de olhos brancos.

24. Uma mosca *Phaenicia sericata* normalmente tem seis cromossomos: dois metacêntricos e quatro acrocêntricos. Um geneticista examina os cromossomos dessas moscas de aparência estranha e descobre que ela tem apenas cinco cromossomos; três são metacêntricos e dois são acrocêntricos. Explique como essa alteração no número de cromossomos poderia ter ocorrido.

*25. Um cromossomo do tipo selvagem tem os segmentos a seguir:

A B C • D E F G H I

Um indivíduo é heterozigoto para as mutações cromossômicas a seguir. Para cada mutação, esboce como os cromossomos do tipo selvagem e mutante seriam pareados na prófase I da meiose, mostrando todas as fitas de cromossomos.

a. A B C • D E F D E F G H I
b. A B C • D H I
c. A B C • D G F E H I
d. A B E D • C F G H I

26. Para os cromossomos apresentados na Figura 8.10, desenhe as cromátides que surgiriam a partir do *crossing over* duplo de fita dupla: um *crossing over* entre C e D e outro entre D e E.

ANÁLISE DE DADOS

*27. Como foi discutido neste capítulo, o *crossing over* em uma inversão pericêntrica produz cromossomos com cópias extras de alguns genes e nenhuma cópia de outros genes. A fertilização dos gametas contendo tal duplicação ou cromossomos deficientes resulta em crianças com síndrome caracterizada por atraso do desenvolvimento, déficit intelectual, desenvolvimento anormal dos sistemas orgânicos e morte precoce. Maarit Jaarola *et al*. examinaram espermatozoides de um homem heterozigoto para uma inversão pericêntrica no cromossomo 8 e determinaram que o *crossing over* ocorria na inversão pericêntrica em 26% das divisões meióticas (M. Jaarola, R. H. Martin, and T. Ashley. 1998. *American Journal of Human Genetics* 63:218-224).

Considere que você é um conselheiro genético, e um casal o procurou para aconselhamento. Ambos são normais do ponto de vista fenotípico, mas a mulher é heterozigota para uma inversão pericêntrica no cromossomo 8. O cariótipo do homem é normal. Qual é a probabilidade de esse casal ter uma criança com uma síndrome debilitante como resultado do *crossing over* na inversão pericêntrica?

*28. Um indivíduo heterozigoto para uma translocação recíproca tem os cromossomos a seguir:

A B • C D E F G
A B • C D V W X
R S • T U E F G
R S • T U V W X

a. Desenhe o rearranjo do pareamento desses cromossomos na prófase I da meiose.
b. Desenhe os padrões de segregação alternada, adjacente-1, adjacente-2 na anáfase I da meiose.
c. Indique os produtos resultantes da segregação alternada, adjacente-1 e adjacente-2.

Seção 8.3

*29. O daltonismo é um distúrbio recessivo ligado ao cromossomo X. Um homem jovem com um cariótipo 47,XXY (síndrome de Klinefelter) é daltônico. Seu irmão 46,XY também é daltônico. Seus pais têm visão colorida normal. Onde ocorreu a não disjunção que deu origem ao homem com a síndrome de Klinefelter? Considere que não ocorre *crossing over* na prófase I da meiose.

ANÁLISE DE DADOS

30. A epidermólise bolhosa juncional (EBJ) é um distúrbio cutâneo grave que resulta em bolhas em todo o corpo. O distúrbio é causado por mutações autossômicas recessivas em qualquer um dos três *loci* que ajudam a codificar a laminina 5, o principal componente na membrana basal da derme-epiderme. Leena Pulkkinen *et al*. descobriram um recém-nascido do sexo masculino com EBJ que morreu com 2 meses de vida (L. Pulkkinen *et al*. 1997. *American Journal of Human Genetics* 61:611-619); a criança tinha pais saudáveis sem parentesco. A análise cromossômica revelou que o recém-nascido tinha 46 cromossomos com aparência normal. A análise de DNA evidenciou que a mãe era heterozigota para o alelo que causava EBJ no *locus LAMB3*, que está no cromossomo 1. O pai tinha dois alelos normais nesse *locus*. O *fingerprinting* do DNA demonstrou que o suposto pai tinha, de fato, concebido a criança.
a. Considerando que não houve mutações nessa família, explique a ocorrência de uma doença recessiva autossômica na criança quando a mãe é heterozigota e o pai é homozigoto normal.
b. Como você poderia comprovar sua explicação? Considere que existem vários marcadores genéticos para cada cromossomo.

31. Algumas pessoas com síndrome de Turner são mosaicos de 45,X/46,XY. Explique como esse mosaicismo poderia surgir.

*32. Bill e Betty tiveram duas crianças com síndrome de Down. O irmão de Bill tinha síndrome de Down, e sua irmã teve duas crianças com síndrome de Down. Com base nessas observações, indique qual das afirmações a seguir é mais provavelmente correta e qual é mais provavelmente incorreta. Explique seu raciocínio.
a. Bill tem 47 cromossomos.
b. Betty tem 47 cromossomos.

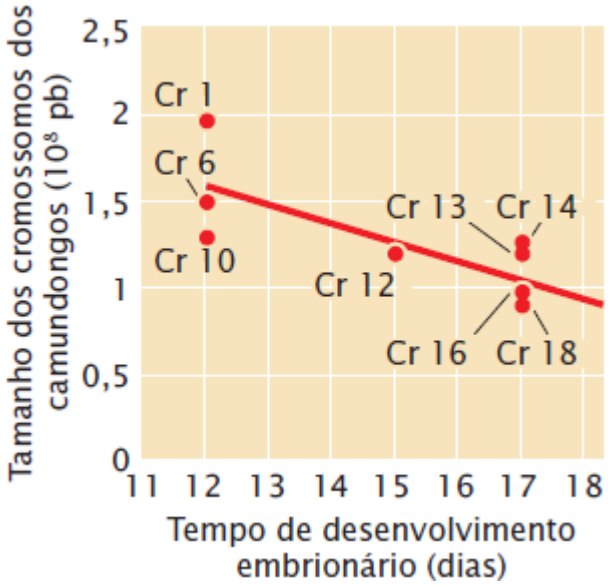
c. As crianças de Bill e Betty têm 47 cromossomos cada uma.
d. A irmã de Bill tem 45 cromossomos.
e. Bill tem 46 cromossomos.
f. Betty tem 45 cromossomos.
g. O irmão de Bill tem 45 cromossomos.

33. Nos mamíferos, os aneuploides dos cromossomos sexuais são mais comuns que os aneuploides autossômicos, mas, no peixe, os aneuploides dos cromossomos sexuais e aneuploides autossômicos são encontrados na mesma frequência. Dê uma possível explicação para essas diferenças nos mamíferos e peixes. (Dica: pense em como as aneuploidias dos cromossomos sexuais são mais comuns que as aneuploidias nos mamíferos.)

*34. Um jovem casal está planejando ter filhos. Ao saber que há um substancial número de natimortos, abortos e problemas de fertilidade na família do marido, eles procuraram um conselheiro genético. A análise cromossômica revela que, enquanto a mulher tem um cariótipo normal, o homem tem 45 cromossomos e é portador de translocação robertsoniana entre os cromossomos 22 e 13.
a. Liste todos os tipos diferentes de gametas que poderiam ser produzidos pelo homem.
b. Que tipos de zigotos se desenvolvem quando cada um dos gametas produzidos pelo homem se une com um gameta normal produzido pela mulher?
c. Se as trissomias e monossomias transmitidas pelos cromossomos 13 e 22 são letais, aproximadamente que proporção de descendentes sobreviventes provavelmente seria portadora da translocação?

ANÁLISE DE DADOS

35. Usando técnicas de procriação, Andrei Dyban e V. S. Baranov (*Cytogenetics of Mammalian Embryonic Development.* Oxford: Oxford University Press, Clarendon Press; New York: Oxford University Press, 1987) criaram camundongos que eram trissômicos para cada um dos diferentes cromossomos de camundongos. Eles descobriram que apenas os animais com trissomia do 19 se desenvolviam. Os camundongos trissômicos para todos os outros cromossomos morriam durante o desenvolvimento. Para alguns desses trissômicos, eles compararam a extensão do desenvolvimento (número de dias após a concepção antes de o embrião morrer) como uma função do tamanho do cromossomo dos camundongos que era encontrado em três cópias (ver o gráfico adjacente). Resuma as descobertas como estão apresentadas nesse gráfico e forneça uma possível explicação para os resultados.



(E. Torres, B. R. Williams, and A. Amon. 2008. Genetics 179:737-746, Fig. 2B.)

Seção 8.4

36. A espécie I tem $2n = 16$ cromossomos. Quantos cromossomos serão encontrados em cada célula em cada um dos mutantes a seguir nessa espécie?
a. Monossômico.
b. Autotriploide.
c. Autotetraploide.
d. Trissômico.
e. Monossômico duplo.

f. Nulissômico.
g. Autopentaploide.
h. Tetrassômico.

37. A espécie I é diploide (2*n* = 8) com os cromossomos AABBCCDD; a espécie relacionada II é diploide (2*n* = 8) com os cromossomos MMNNOOPP. Que tipos de mutações cromossômicas os organismos com os seguintes conjuntos de cromossomos apresentam?
    a. AAABBCCDD.
    b. MMNNOOOOPP.
    c. AABBCDD.
    d. AAABBBCCCDDD.
    e. AAABBCCDDD.
    f. AABBDD.
    g. AABBCCDDMMNNOOPP.
    h. AABBCCDDMNOP.

*38. A espécie I tem 2*n* = 8 e a espécie II tem 2*n* = 14 cromossomos. Quais seriam os números de cromossomos esperados em organismos individuais com as seguintes mutações cromossômicas? Indique todas as possíveis respostas.
    a. Alotriploidia incluindo as espécies I e II.
    b. Autotetraploidia na espécie II.
    c. Trissomia na espécie I.
    d. Monossomia na espécie II.
    e. Tetrassomia na espécie I.
    f. Alotetraploidia incluindo as espécies I e II.

39. Suponha que a espécie I na Figura 8.26 tenha 2*n* = 10 e a espécie II na figura tenha 2*n* = 12. Quantos cromossomos existiriam no alotetraploide no final da figura?

40. Considere uma célula diploide com 2*n* = 4 cromossomos – um par de cromossomos metacêntricos e um par de cromossomos acrocêntricos. Suponha que essa célula sofra não disjunção, dando origem a uma célula autotriploide (3*n*). A célula triploide então sofre meiose. Desenhe os diferentes tipos de gameta que podem surgir a partir da meiose na célula triploide, mostrando os cromossomos presentes em cada tipo. Use uma cor diferente para cada cromossomo metacêntrico a fim de diferenciar os cromossomos metacêntricos e acrocêntricos; da mesma forma, use uma cor diferente para cada cromossomo acrocêntrico. (Dica: ver Figura 8.25.)

41. Considere que a célula autotriploide na Figura 8.25 tem 3*n* = 30 cromossomos. Para cada um dos gametas produzidos por essa célula, indique o número de cromossomos do zigoto resultante se o gameta se unir com um gameta haploide normal.

ANÁLISE DE DADOS

42. *Nicotiana glutinosa* (2*n* = 24) e *N. tabacum* (2*n* = 48) são duas plantas muito próximas que podem ser intercruzadas, mas as plantas híbridas F_1 resultantes, em geral, são estéreis. Em 1925, Roy Clausen e Thomas Goodspeed cruzaram *N. glutinosa* e *N. tabacum* e obtiveram uma planta F_1 fértil (R. E. Clausen and T. H. Goodspeed. 1925. *Genetics* 10:278-284). Eles conseguiram autopolinizar as flores dessa planta para produzir uma geração F_2. Surpreendentemente, as plantas F_2 eram férteis e produziam sementes viáveis. Quando Clausen e Goodspeed examinaram os cromossomos das plantas F_2, observaram 36 pares de cromossomos na metáfase I e 36 cromossomos individuais na metáfase II. Explique a origem das plantas F_2 obtidas por Clausen e Goodspeed e os números de cromossomos observados.

43. Qual seria o número de cromossomos dos descendentes dos cruzamentos a seguir no trigo (ver Figura 8.29)? Que tipo de poliploidia (alotriploidia, alotetraploidia etc.) surgiria a partir de cada cruzamento?
    a. Trigo Einkorn e trigo emmer (farro).
    b. Trigo de pão e trigo emmer (farro).
    c. Trigo Einkorn e trigo de pão.

ANÁLISE DE DADOS

44. Karl e Hally Sax cruzaram *Aegilops cylindrica* (2*n* = 28), uma grama selvagem encontrada na região do Mediterrâneo, com *Triticum vulgare* (2*n* = 42), um tipo de trigo (K. Sax and H. J. Sax. 1924. *Genetics* 9:454-464). As plantas F_1 resultantes desse cruzamento tinham 35 cromossomos. O exame da metáfase I das plantas F_1 revelou 7 pares de cromossomos (bivalentes) e 21 pares de cromossomos não pareados (univalentes).
    a. Se os cromossomos não pareados se separam de forma aleatória, que possíveis números de cromossomos aparecerão nos gametas das plantas F_1?

b. O que a aparência dos bivalentes nos híbridos de F_1 sugere sobre a origem do trigo *Triticum vulgare*?

Questões desafiadoras

Seção 8.3

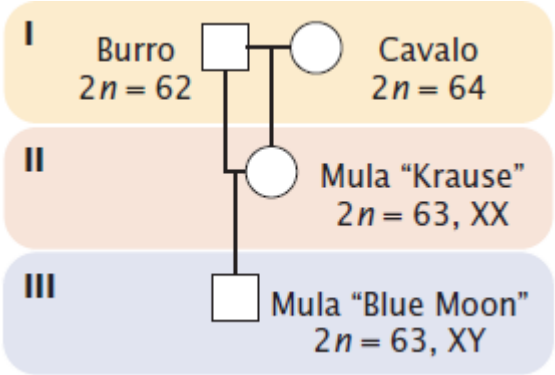
45. O daltonismo é um distúrbio recessivo ligado ao cromossomo X. Jill tem a visão normal, mas seu pai é daltônico. Ela se casou com Tom, que também tem visão normal. Jill e Tom têm uma filha daltônica com síndrome de Turner.
    a. Como sua filha herdou o daltonismo?
    b. A filha herdou o cromossomo X de Jill ou de Tom?

ANÁLISE DE DADOS

46. Os descendentes dos tomates triploides têm partes de um cromossomo extra, além do complemento normal de 24 cromossomos (J. W. Lesley and M. M. Lesley. 1929. *Genetics* 14:321-336). Os mutantes com uma parte de um cromossomo extra são chamados de secundários. James e Margaret Lesley observaram que os secundários surgiam a partir de genitores triploides ($3n$), trissômicos ($3n + 1$) e duplo-trissômicos ($3n + 1 + 1$), mas nunca de diploides ($2n$). Indique uma ou mais possíveis explicações para os secundários que surgem a partir de genitores com cromossomos não pareados, mas não de genitores que são diploides normais.

ANÁLISE DE DADOS

47. As mulas resultam do cruzamento entre um cavalo ($2n = 64$) e um burro ($2n = 62$), têm 63 cromossomos e quase todas são estéreis. Entretanto, no verão de 1985, uma mula-fêmea chamada Krause que pastou com um burro-macho deu à luz um filhote (O. A. Ryder *et al*. 1985. *Journal of Heredity* 76:379-381). Os testes sanguíneos indicaram que o filhote macho, chamado de Blue Moon, era o descendente de Krause e que ela era de fato uma mula. Blue Moon e Krause eram oriundos do mesmo burro (ver ilustração). O filhote, como sua mãe, tinha 63 cromossomos – metade deles cromossomos de cavalo, e a outra metade cromossomos de burro. As análises dos marcadores genéticos mostraram que, extraordinariamente, Blue Moon tinha herdado um conjunto completo de cromossomos equinos de sua mãe, em vez da mistura aleatória dos cromossomos de cavalo e burro que seriam esperados com a meiose normal. Assim, Blue Moon e Krause não eram apenas mãe e filho, mas também irmãos.
    a. Com o uso de um diagrama, mostre como, se Blue Moon herdou apenas cromossomos equinos de sua mãe, ele e Krause são mãe e filho assim como irmãos.
    b. Embora sejam raros, foram relatados outros casos de mulas férteis dando à luz. Nesses casos, quando uma mula-fêmea cruza com um cavalo-macho, os descendentes têm aparência de cavalo, mas, quando uma mula-fêmea cruza com um burro-macho, os descendentes têm aparência de mula. Essa observação é consistente com a ideia de que os descendentes da mula-fêmea fértil herdaram apenas um único conjunto de cromossomos equinos da mula genitora? Explique seu raciocínio.
    c. Você pode sugerir um possível mecanismo para como os descendentes da mula-fêmea fértil podem passar um conjunto completo de cromossomos equinos para seus descendentes?



Seção 8.4

48. Os seres humanos e muitos outros organismos complexos são diploides, com dois conjuntos de genes, um herdado da genitora e outro do genitor. Entretanto, vários organismos eucariotos passam a maior parte do seu ciclo de vida no estado haploide. Muitos desses eucariotos, como *Neurospora* e levedura, ainda sofrem meiose e reprodução sexuada, mas a maioria das células que constitui o organismo é haploide.

Considerando que os organismos haploides são totalmente capazes da reprodução sexuada e de gerar variação genética, por que os eucariotos mais complexos são diploides? Em outras palavras, qual seria a vantagem evolutiva de ter um estado diploide em vez do estado haploide? E por que alguns organismos, como *Neurospora* e levedura, poderiam existir como haploides?

9 Sistemas Genéticos das Bactérias e dos Vírus

A vida no mundo das bactérias

Nós, seres humanos, gostamos de pensar que controlamos o mundo, mas, se comparados com as bactérias, somos minoria. As bactérias evoluíram há 3,5 bilhões de anos, 2 bilhões de anos antes de os primeiros eucariotos surgirem (alguns indícios sugerem que as bactérias evoluíram até antes disso). Hoje em dia, as bactérias são encontradas em qualquer ambiente concebível, incluindo fontes termais, lagos muito salinos e abaixo de mais de 3 km de gelo na Antártica. São encontradas no topo do monte Everest e nas partes mais profundas dos oceanos. Também estão dentro de *nós* – em números alarmantes! No intestino humano há aproximadamente 10 trilhões de bactérias, dez vezes o número total de células do corpo humano. Ninguém sabe quantas bactérias existem no mundo, mas uma análise feita por cientistas em 1998 estimou que o número total de bactérias vivas na Terra exceda 5 milhões de trilhões de trilhões (5×10^{30}).

As bactérias não são apenas numericamente superiores; elas também constituem a maioria da diversidade de vida. O número total de espécies de bactérias descritas é menos de 10.000, comparado com cerca de 1,4 milhão de plantas, animais, fungos e eucariotos unicelulares. No entanto, o número de espécies de bactérias descritas está muito abaixo da verdadeira diversidade microbiana.

Tipicamente, espécies de bactérias são descritas apenas após terem sido cultivadas e estudadas em laboratório. Como apenas algumas delas podem ser cultivadas em laboratório, por muitos anos era impossível identificar e estudar a maioria das bactérias. Então, na década de 1970, surgiram as técnicas moleculares para análise de DNA, que trouxeram uma imagem totalmente nova sobre a diversidade microbiana. Essas técnicas revelaram diversos fatores importantes sobre bactérias. Primeiro, muitas das relações entre as bactérias que os microbiologistas solucionavam com base nos traços físicos e bioquímicos mostraram-se incorretas. As bactérias que pareciam próximas, na verdade, eram bem diferentes do ponto de vista genético. Segundo, a análise molecular mostrou que os membros de um grupo de micróbios – atualmente chamados de archaea – eram diferentes das bactérias (as eubactérias) assim como as bactérias são diferentes dos eucariotos. Terceiro, a análise molecular revelou que o número de diferentes tipos de bactérias é impressionante.

Em 2007, Luiz Roesch e colaboradores começaram a determinar exatamente quantos tipos de bactérias existem em um grama de solo. Eles obtiveram amostras de solo de quatro locais – Brasil, Flórida, Illinois e Canadá – e, a partir delas, extraíram e purificaram o DNA bacteriano. Tendo como ponto de partida esse DNA, eles determinaram as sequências de um gene existente em todas as bactérias, o 16S rRNA. Cada espécie diferente de bactéria tem uma sequência única do gene 16S rRNA, então eles podiam determinar quantas espécies de bactérias existiam em cada amostra de solo ao contar o número de diferentes sequências de DNA.

Os resultados de Roesch foram impressionantes. O número de espécies eubacterianas diferentes em cada grama de solo variou de 26.140 amostras do Brasil a 53.533 do Canadá. Foram detectadas muitas bactérias incomuns que pareciam ser diferentes de todos os grupos de bactérias já descritos. Outro achado interessante foi que o solo dos campos de agricultura abrigava muito menos espécies do que as amostras de solo de florestas.

Esse estudo e outros mais demonstraram que a diversidade bacteriana é muito maior que a dos organismos multicelulares, e, sem dúvida, vários grupos de bactérias não foram descobertos ainda. Querendo ou não, na verdade, nós é que vivemos no mundo das bactérias.

Neste capítulo, examinaremos algumas das propriedades genéticas das bactérias e dos vírus e os mecanismos que eles usam para trocar e recombinar seus genes. Desde 1940, os sistemas genéticos das bactérias e vírus contribuem para a descoberta de muitos conceitos importantes na genética. O estudo da genética molecular se concentra quase totalmente em seus genes; atualmente, as bactérias e os vírus ainda são ferramentas essenciais para sondar a natureza dos genes em organismos mais complexos, em parte porque eles têm várias características que os tornam adequados para estudos genéticos (**Quadro 9.1**).

Os sistemas genéticos das bactérias e vírus também são objeto de estudo, pois esses organismos têm papel fundamental em

nossa sociedade. As bactérias são encontradas naturalmente na boca, no intestino e na pele, onde são essenciais para o ser humano e a ecologia. Elas estão aptas para produzir várias substâncias importantes para a Economia e têm grande importância no campo da Medicina, sendo responsáveis por várias doenças. Neste capítulo vamos estudar os vários aspectos únicos dos sistemas genéticos das bactérias e dos vírus. Processos importantes de transferência e recombinação de genes estão descritos neste capítulo e veremos como eles podem ser usados para mapear os genes bacterianos e virais. **Resolva o Problema 17**

9.1 A análise genética das bactérias exige métodos especiais

A hereditariedade nas bactérias é muito semelhante à dos organismos mais complexos, mas o genoma haploide das bactérias e seu pequeno tamanho (o que dificulta a observação dos seus fenótipos) exigem abordagens e métodos diferentes.

Diversidade bacteriana

Os procariotos são organismos unicelulares que não têm membrana nuclear nem organelas celulares ligadas a membranas. Por muitos anos, os biólogos consideravam que todos os procariotos eram próximos, mas as informações das sequências de DNA fornecem evidência clara de que eles são divididos em pelo menos dois grupos diferentes: archaea e eubactérias. As archaea são um grupo de procariotos diversos encontrados com frequência em ambientes extremos, como fontes termais e no fundo dos oceanos. As eubactérias são os procariotos remanescentes, incluindo a maioria das bactérias familiares. Embora sejam superficialmente semelhantes em sua estrutura celular, as eubactérias e as archaea são diferentes em sua composição genética, e as diferenças entre elas são maiores do que as diferenças entre as eubactérias e os eucariotos. Na verdade, as archaea são *mais* semelhantes aos eucariotos que as eubactérias em várias características moleculares e processos genéticos. O termo bactéria será usado para as eubactérias neste livro.

Quadro 9.1 Vantagens do uso de bactérias e vírus para estudos genéticos.

1. A reprodução é rápida.
2. São produzidos muitos descendentes.
3. O genoma haploide permite que todas as mutações sejam expressas de forma direta.
4. A reprodução assexuada simplifica o isolamento de cepas geneticamente puras.
5. O crescimento no laboratório é simples e requer pouco espaço.
6. Os genomas são pequenos.
7. Existem técnicas disponíveis para isolamento e manipulação de seus genes.
8. São importantes para a Medicina.
9. Podem ser alterados por meio de engenharia genética para produzir substâncias com valor comercial.

As bactérias são muito diferentes e têm vários formatos e tamanhos. Algumas são bacilares, enquanto outras são esféricas. A maioria é menor que as células eucarióticas, mas pelo menos uma espécie isolada do intestino de peixe tem quase 1 mm de comprimento e pode ser observada a olho nu. Algumas bactérias são fotossintéticas. Outras produzem talos e esporos, lembrando um fungo.

As bactérias já foram consideradas organismos simples porque não têm boa parte da complexidade celular dos eucariotos. Entretanto, dados recentes indicam várias semelhanças e paralelos nas estruturas bacteriana e eucariótica. Por exemplo, uma proteína bacteriana chamada FtsZ, que tem papel importante na divisão celular, é estruturalmente semelhante às proteínas tubulinas dos eucariotos, que são subunidades de microtúbulos e ajudam a separar os cromossomos na mitose e meiose (ver *Capítulo 2*). Como os eucariotos, as bactérias têm proteínas que ajudam a condensar o DNA. Outras proteínas bacterianas funcionam como proteínas citoesqueléticas como nos eucariotos, ajudando a dar o formato e a estrutura das bactérias. E, embora as bactérias não sofram mitose nem meiose, a replicação do cromossomo bacteriano precede a fissão binária, e existem processos bacterianos que garantem que uma cópia do cromossomo seja alocada em cada célula.

Técnicas para o estudo das bactérias

A cultura e o estudo das bactérias requerem técnicas especiais. Os microbiologistas definiram as necessidades nutricionais de

várias bactérias e desenvolveram meios de cultura para elas crescerem no laboratório. Em geral, os meios de cultura têm uma fonte de carbono, elementos essenciais como nitrogênio e fósforo, algumas vitaminas e outros íons e nutrientes necessários. As bactérias do tipo selvagem, ou **prototróficas**, conseguem usar esses ingredientes simples para sintetizar todos os compostos necessários ao crescimento e à reprodução. Um meio com apenas os nutrientes exigidos pelas bactérias prototróficas é chamado de **meio de cultura mínimo**.

As cepas mutantes conhecidas como **auxotróficas** não têm uma ou mais enzimas necessárias para sintetizar as moléculas essenciais e crescem apenas no meio suplementado com essas moléculas essenciais. Por exemplo, as cepas auxotróficas incapazes de sintetizar o aminoácido leucina não crescem em meio mínimo, mas *crescerão* no meio ao qual a leucina foi adicionada. O **meio de cultura completo** tem todas as substâncias, como o aminoácido leucina, necessárias ao crescimento e à reprodução das bactérias.

As culturas de bactérias crescem em tubos de teste com meio líquido estéril (**Figura 9.1 A**). Algumas bactérias são adicionadas ao tubo e, então, crescem e se dividem até usarem todos os nutrientes ou – mais comumente – até a concentração de seus dejetos tornar-se tóxica. As bactérias também crescem em placas Petri (**Figura 9.1 B**). O meio de crescimento suspenso em ágar é transferido para a metade inferior da placa de Petri, fornecendo uma base sólida, semelhante a gel, para o crescimento bacteriano. Em um processo chamado plaqueamento, uma solução diluída da bactéria é espalhada sobre a superfície de uma placa Petri preenchida com ágar. À medida que as bactérias crescem e se dividem, elas originam uma massa visível de células geneticamente idênticas (uma **colônia**). As cepas geneticamente puras das bactérias podem ser isoladas ao coletar as bactérias de uma única colônia e transferi-las para um novo tubo de teste ou placa Petri. A principal vantagem desse método é que ele possibilita isolar e contar as bactérias, que individualmente são pequenas demais para serem observadas sem um microscópio.

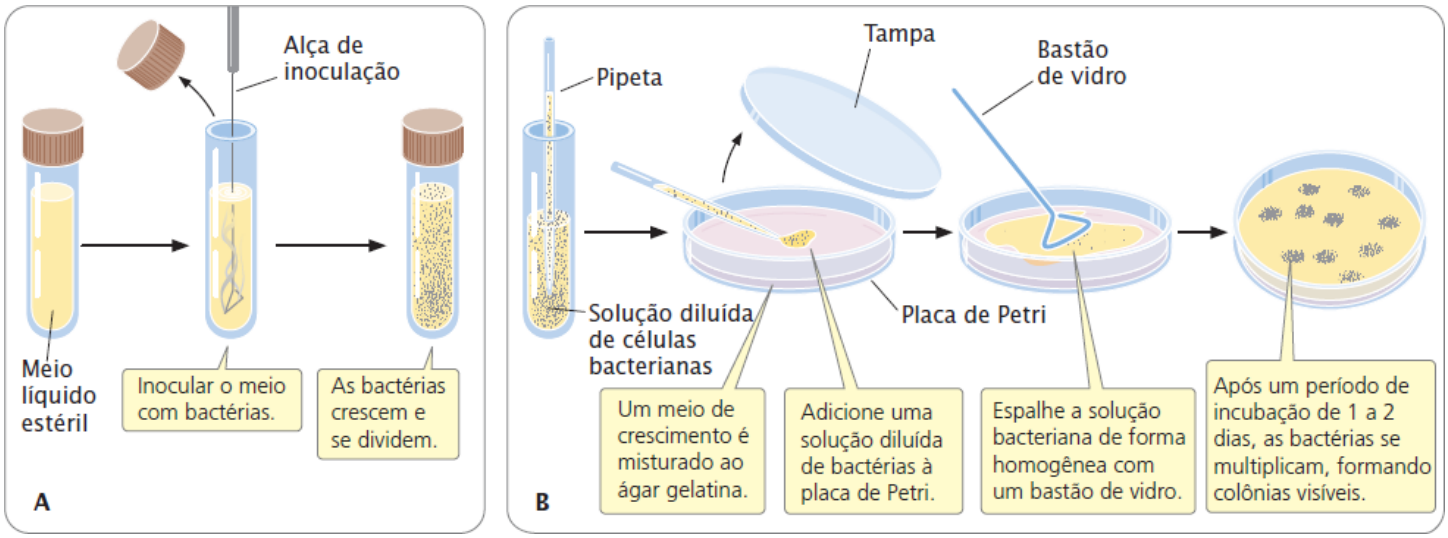


Figura 9.1 As bactérias podem crescer em um meio (A) líquido ou (B) sólido.

Frequentemente, os microbiologistas estudam fenótipos que afetam a aparência da colônia (**Figura 9.2**) ou podem ser detectados por testes químicos simples. É comum estudar os fenótipos dos auxotróficos. Suponha que queremos detectar auxotróficos que não conseguem sintetizar leucinas (*leu*$^-$ mutantes). Primeiro, espalhamos as bactérias em uma placa de Petri com um meio rico em leucina; os prototróficos que têm um alelo *leu*$^+$ e os auxotróficos que têm o alelo *leu*$^-$ crescem nesse meio (**Figura 9.3**). A seguir, por meio da técnica chamada *replica plating* (plaqueamento em réplica), transferimos algumas células de cada uma das colônias da placa original para duas novas placas de reprodução: uma com o meio rico em leucina e outra com um meio sem leucina. Um meio sem um nutriente essencial, como o meio sem leucina, é chamado de meio seletivo. As bactérias *leu*$^+$ crescerão em ambos os meios, mas as mutantes *leu*$^-$ crescerão apenas no meio seletivo enriquecido com leucina, porque elas não podem sintetizar sua própria leucina. Qualquer colônia que cresce no meio com leucina, mas que não cresce no meio sem leucina, consiste em bactérias *leu*$^-$. Os auxótrofos que crescem no meio enriquecido podem, então, ser cultivados para mais estudo.

O genoma bacteriano

A maioria dos genomas bacterianos tem um cromossomo circular com uma única molécula de DNA com milhares de pares de bases (pb) de comprimento. Por exemplo, o genoma da *E. coli* tem aproximadamente 4,6 milhões de pares de base de DNA. Entretanto, algumas bactérias têm múltiplos cromossomos. Por exemplo, *Vibrio cholerae*, que causa a cólera, tem dois cromossomos circulares, e *Rhizobium meliloti* tem três cromossomos. Existem até algumas bactérias com cromossomos lineares. Muitos cromossomos bacterianos estão organizados de forma eficiente. Por exemplo, mais de 90% do DNA na *E. coli*

codificam proteínas. Por outro lado, apenas 1% do DNA humano codifica proteínas.

Plasmídios

Além de ter um cromossomo, muitas bactérias têm **plasmídios** – pequenas moléculas circulares de DNA. Pode haver alguns plasmídios em muitas cópias por célula, ou em apenas uma ou duas cópias. Em geral, os plasmídios carreiam genes não essenciais ao funcionamento da bactéria, mas importantes no ciclo de vida e no crescimento das suas bactérias hospedeiras. Existem muitos tipos diferentes de plasmídios; estima-se que a *E. coli* sozinha tenha mais de 270 plasmídios naturais. Alguns promovem o cruzamento entre bactérias, outros têm genes que matam outras bactérias. Os plasmídios são muito usados na engenharia genética (ver *Capítulo 19*), e alguns participam na disseminação da resistência a antibióticos entre as bactérias.

Figura 9.2 As bactérias têm vários fenótipos. A. *Serratia marcesens* com variação de cor. **B.** *Bacillus cereus*. (*Parte A*: Dr. Edward J. Bottone. *Parte B*: Biophoto Associates/Photo Researchers.)

A maioria dos plasmídios é circular e tem milhares de pares de bases de comprimento, embora sejam encontrados plasmídios com centenas de milhares de pares de bases. Cada plasmídio tem uma origem de replicação, uma sequência específica de DNA onde inicia a replicação (ver *Capítulo 2*). A origem permite que o plasmídio se replique independentemente do cromossomo bacteriano (**Figura 9.4**). Os **episssomos** são plasmídios capazes de replicação livre e de se integrar aos cromossomos bacterianos. O **fator F (fertilidade)** da *E. coli* (**Figura 9.5**) é um episssomo que controla o cruzamento e a troca de genes entre as células dessa bactéria, como será discutido na próxima seção.

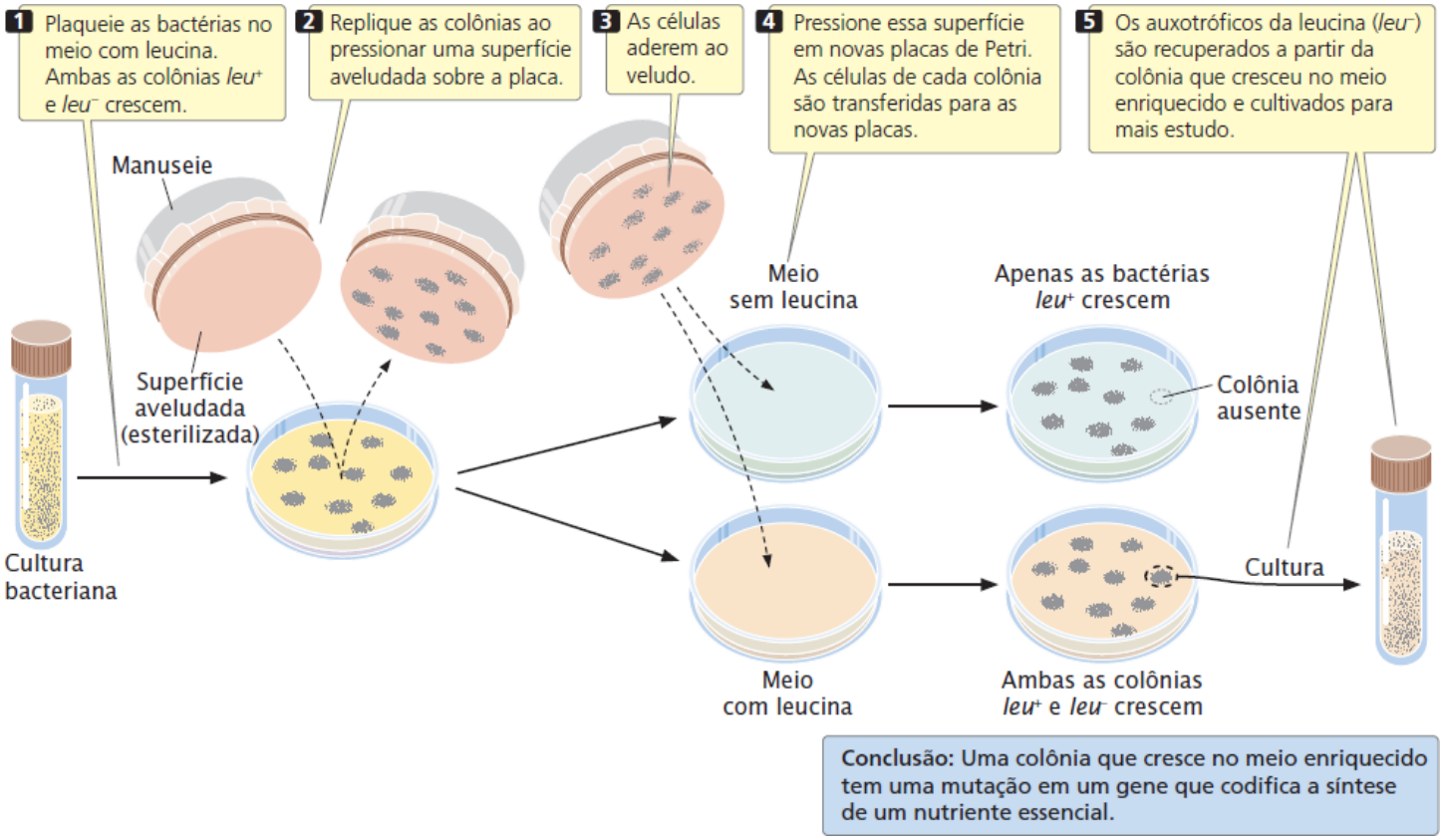


Figura 9.3 As cepas bacterianas mutantes podem ser isoladas com base em suas exigências nutricionais.

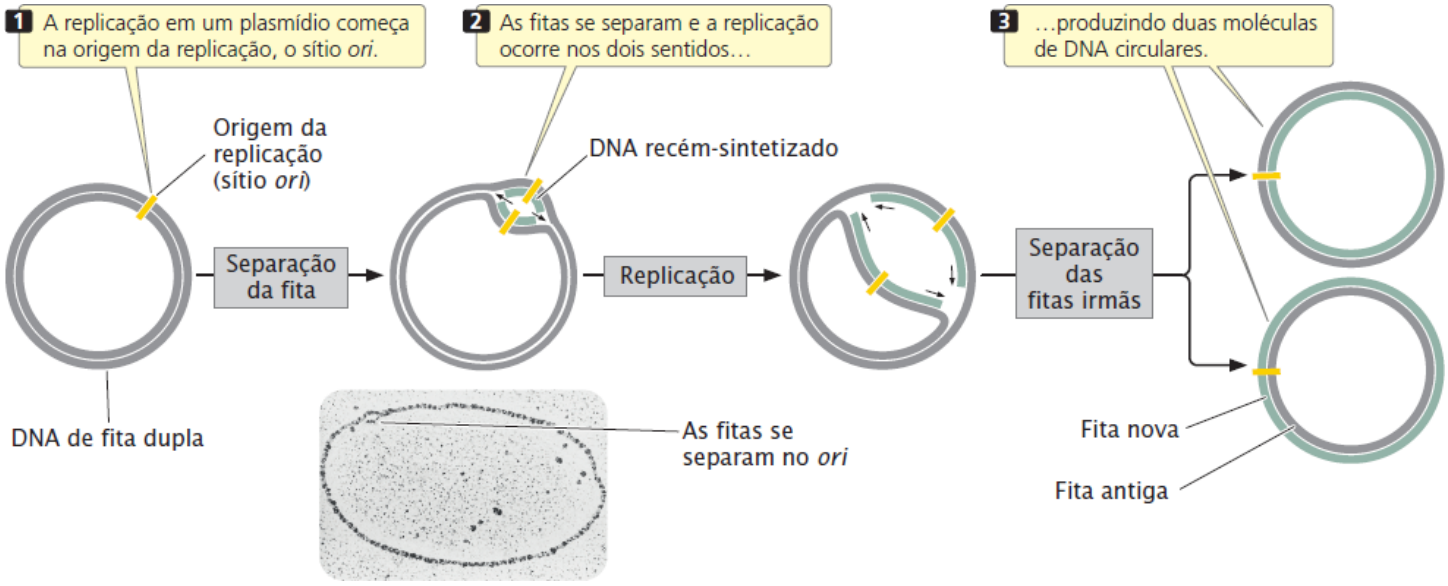


Figura 9.4 Um plasmídio se replica independentemente de seu cromossomo bacteriano. A replicação começa na origem da replicação (*ori*) e continua ao redor do círculo. Nesse diagrama, a replicação está ocorrendo nos dois sentidos; em alguns plasmídios, a replicação ocorre apenas em um sentido. (Fotografia: Biology Pics/Science Source.)

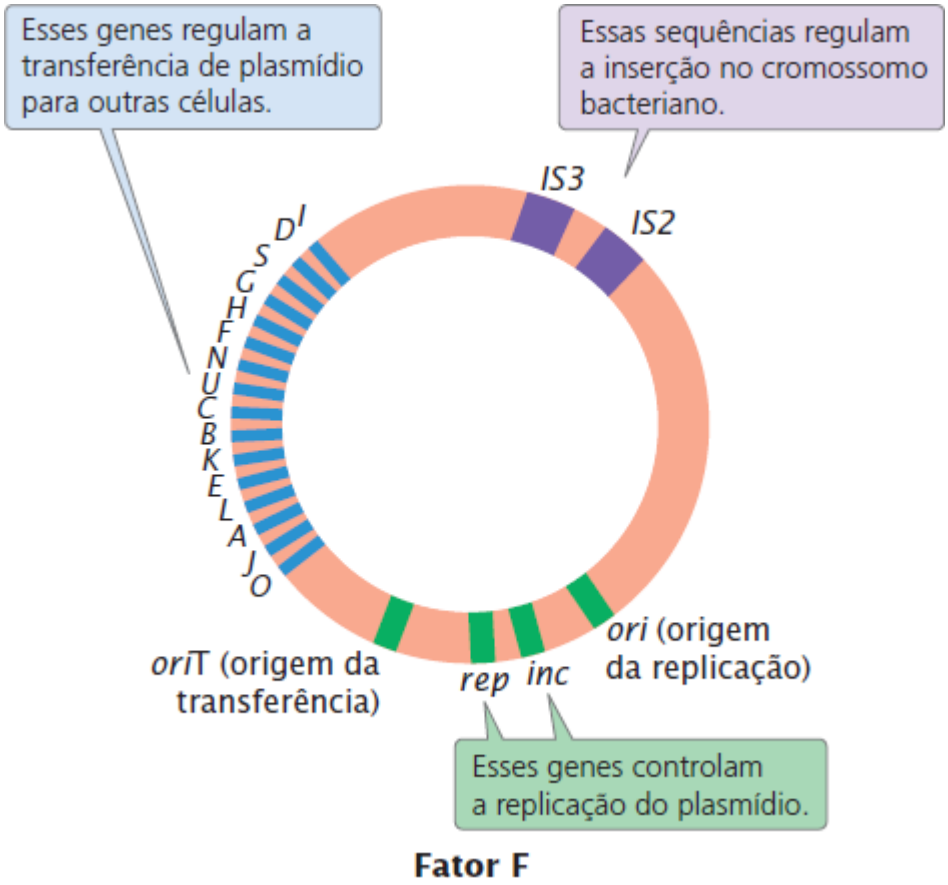


Figura 9.5 O fator F, um episssomo circular da *E. coli*, tem vários genes que regulam a transferência para a célula bacteriana, a replicação e a inserção no cromossomo bacteriano. A replicação se inicia no *ori*. As sequências de inserção (ver *Capítulo 18*) *IS3* e *IS2* controlam a inserção e a retirada do cromossomo bacteriano.

Conceitos

As bactérias podem ser estudadas no laboratório por cultivo em meio líquido ou sólido definido. Um genoma bacteriano típico consiste em um único cromossomo circular com milhares de pares de bases. Alguns genes bacterianos estão nos plasmídios, que são pequenas moléculas circulares de DNA que se replicam independentemente do cromossomo bacteriano.

Checagem dos conceitos 1

O que é verdadeiro sobre os plasmídios?

a. São compostos de RNA.
b. Existem normalmente fora das células bacterianas.
c. Eles têm apenas uma única fita de DNA.
d. Eles têm uma origem de replicação.

9.2 As bactérias trocam os genes por meio de conjugação, transformação e transdução

As bactérias trocam material genético por meio de três mecanismos diferentes, todos exigindo algum tipo de transferência de DNA e recombinação entre o DNA transferido e o cromossomo bacteriano.

1. A **conjugação** ocorre quando o material genético passa diretamente de uma bactéria para outra (**Figura 9.6 A**). Na conjugação, duas bactérias se aproximam e forma-se uma conexão entre elas. Um plasmídio ou uma parte do cromossomo bacteriano passa de uma célula (a doadora) para outra (a receptora). Após a conjugação, ocorre *crossing over* entre as sequências homólogas no DNA transferido e o cromossomo da célula receptora. Na conjugação, o DNA é transferido apenas da doadora para a receptora, sem troca recíproca de material genético.
2. A **transformação** ocorre quando uma bactéria capta o DNA do meio onde está crescendo (**Figura 9.6 B**). Após a transformação, pode ocorrer recombinação entre os genes introduzidos e os genes do cromossomo bacteriano.

3. A **transdução** ocorre quando os vírus bacterianos (bacteriófagos) transportam DNA de uma bactéria para outra (**Figura 9.6 C**). Dentro da bactéria, o DNA recém-introduzido pode sofrer recombinação com o cromossomo bacteriano.

Nem todas as espécies de bactérias exibem os três tipos de transferência genética. A conjugação se dá com maior frequência em algumas espécies do que em outras. A transformação ocorre em uma extensão limitada em muitas espécies de bactérias, mas as técnicas laboratoriais aumentam a taxa de captação de DNA. A maioria dos bacteriófagos tem uma gama de hospedeiros limitada; assim, a transdução ocorre normalmente apenas entre bactérias da mesma espécie ou de espécies próximas.

Esses processos de troca genética nas bactérias são diferentes da reprodução sexuada dos eucariotos de duas formas importantes. Primeiro, a troca de DNA e a reprodução não estão acopladas nas bactérias; elas sofrem reprodução (divisão celular) sem receber nenhum DNA de outra célula. Segundo, o material genético doado que não é recombinado no DNA hospedeiro é degradado, então a célula receptora permanece haploide. Cada tipo de transferência genética pode ser usado para mapear genes, como será discutido nas seções a seguir.

Conceitos

O DNA pode ser transferido entre células bacterianas por meio da conjugação, transformação ou transdução. Cada tipo de transferência genética é um movimento unidirecional da informação genética para a célula receptora, às vezes seguida de recombinação. Esses processos não estão conectados à reprodução celular nas bactérias.

Checagem dos conceitos 2

Qual processo de transferência de DNA nas bactérias precisa de um vírus?

a. Conjugação.
b. Transdução.
c. Transformação.
d. Todas as opções acima.

Conjugação

Em 1946, Joshua Lederberg e Edward Tatum demonstraram que as bactérias podem transferir e recombinar informação genética, possibilitando o uso das bactérias nos estudos genéticos. Durante sua pesquisa, Lederberg e Tatum estudaram as cepas auxotróficas de *E. coli*. A cepa Y10 precisava dos aminoácidos treonina (e era genotipicamente thr^-) e leucina (leu^-) e a vitamina tiamina (thi^-) para o crescimento, mas não precisava da vitamina biotina (bio^+) ou dos aminoácidos fenilalanina (phe^+) e cisteína (cys^+); o genótipo dessa cepa pode ser escrito como $thr^-\ leu^-\ thi^-\ bio^+\ phe^+\ cys^+$. A cepa Y24 tinha o conjunto oposto de alelos: ela precisava de biotina, fenilalanina e cisteína no seu meio, mas não precisava de treonina, leucina ou tiamina; seu genótipo era $thr^+\ leu^+\ thi^+\ bio^-\ phe^-\ cys^-$. Em um experimento, Lederberg e Tatum misturaram bactérias Y10 e Y24 e plaquearam em meio mínimo (**Figura 9.7**). Cada cepa também foi plaqueada separadamente em meio mínimo.

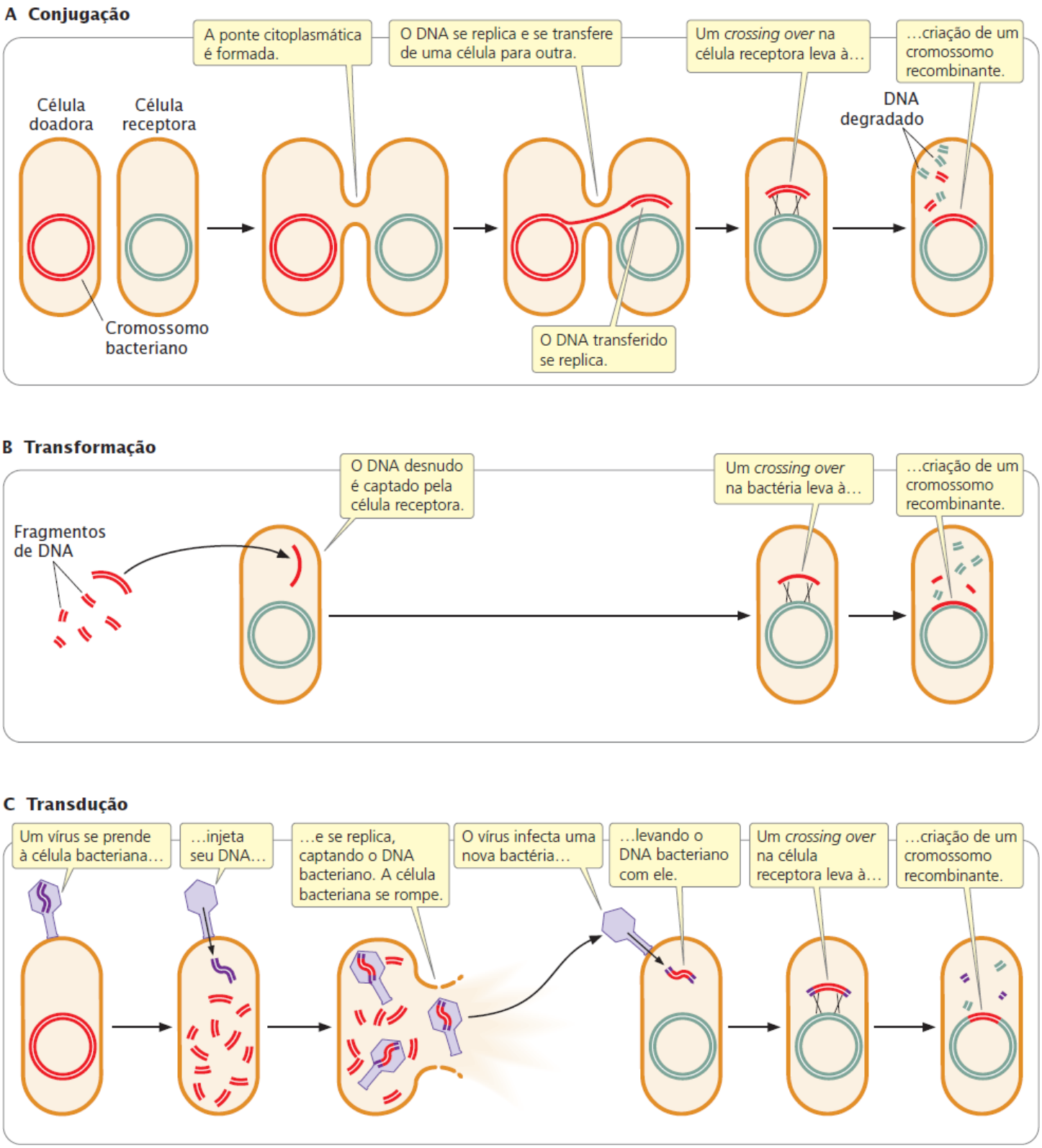


Figura 9.6 Conjugação, transformação e transdução são os três processos da transferência de genes nas bactérias. Para que o DNA transferido seja herdado de forma estável, os três processos precisam que o DNA transferido se recombine com o cromossomo bacteriano.

Sozinhas, as cepas Y10 e Y24 não cresceram em meio mínimo: cada uma delas precisava de nutrientes que o meio não oferecia. A cepa Y10 era incapaz de crescer porque precisava de treonina, leucina e tiamina, que não existiam em meio mínimo; a cepa Y24 era incapaz de crescer, porque precisava de biotina, fenilalanina e cisteína, também ausentes nesse meio. Quando Lederberg e Tatum misturaram as duas cepas, entretanto, algumas colônias cresceram em meio mínimo. Essas bactérias prototróficas têm obrigatoriamente o genótipo thr^+ leu^+ thi^+ bio^+ phe^+ cys^+. De onde veio este genótipo?

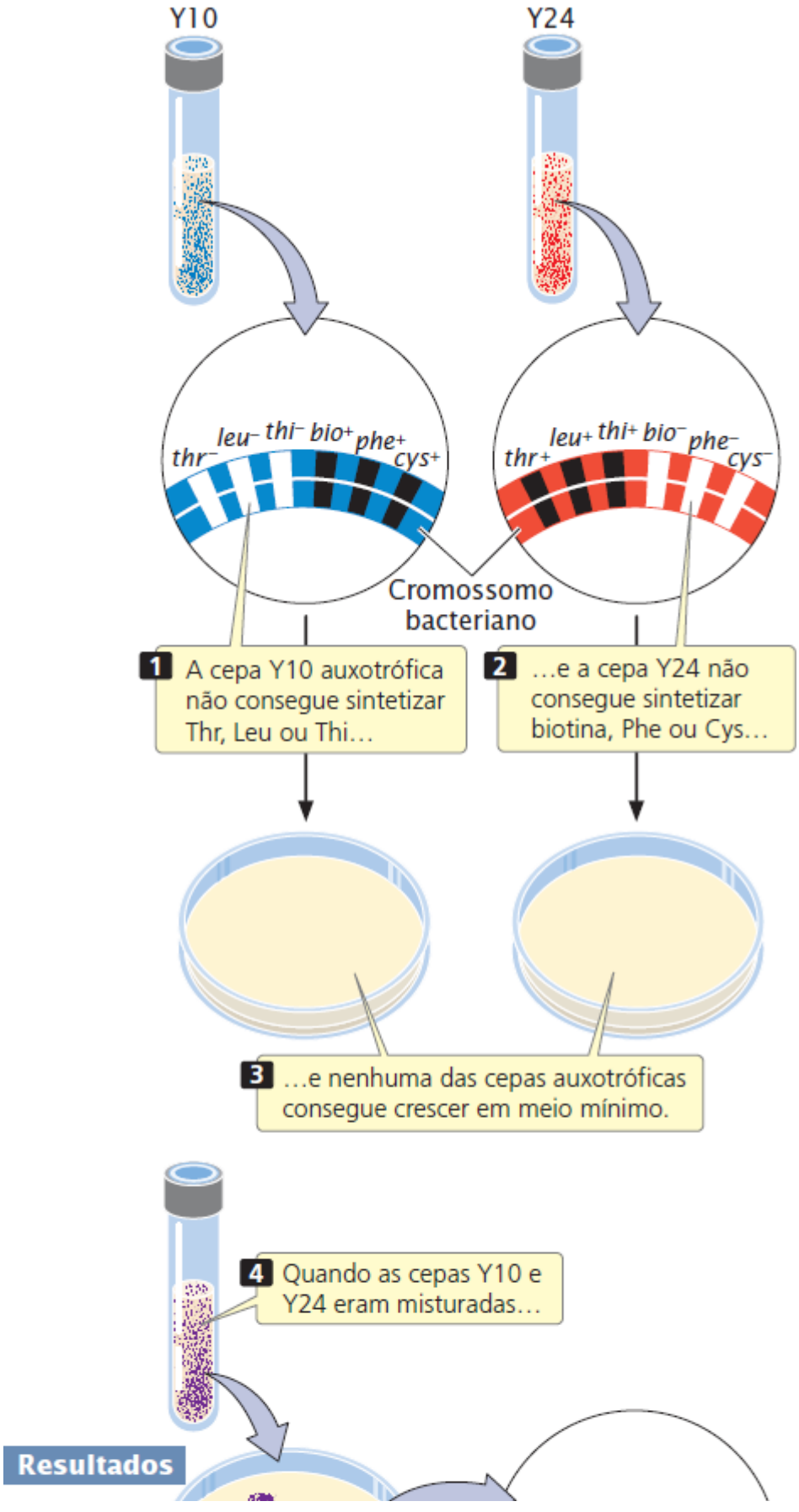
Se as mutações eram responsáveis pelas colônias prototróficas, então algumas delas deveriam ter crescido nas placas contendo Y10 e Y24 sozinhas, o que não aconteceu. Seriam necessárias múltiplas mutações simultâneas ($thr^- \rightarrow thr^+$, $leu^- \rightarrow leu^+$ e $thi^- \rightarrow thi^+$ na cepa Y10 ou $bio^- \rightarrow bio^+$, $phe^- \rightarrow phe^+$ e $cys^- \rightarrow cys^+$ na cepa Y24) para que cada cepa se tornasse prototrófica por meio da mutação, o que era pouco provável. Lederberg e Tatum concluíram que ocorreu algum tipo de

transferência e recombinação genética:

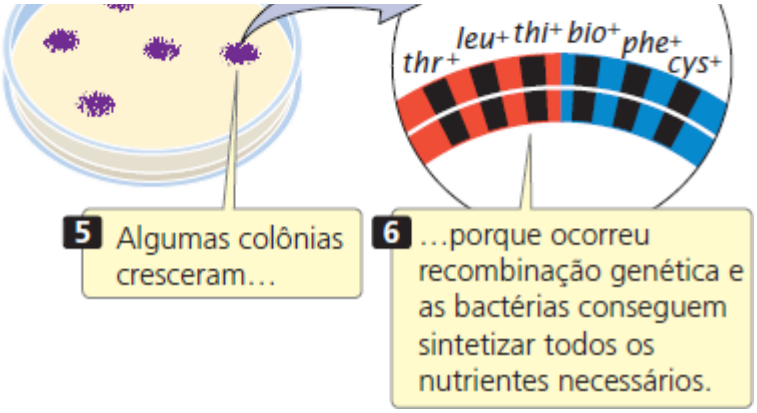
Experimento

Pergunta: As bactérias trocam informações genéticas?

Métodos

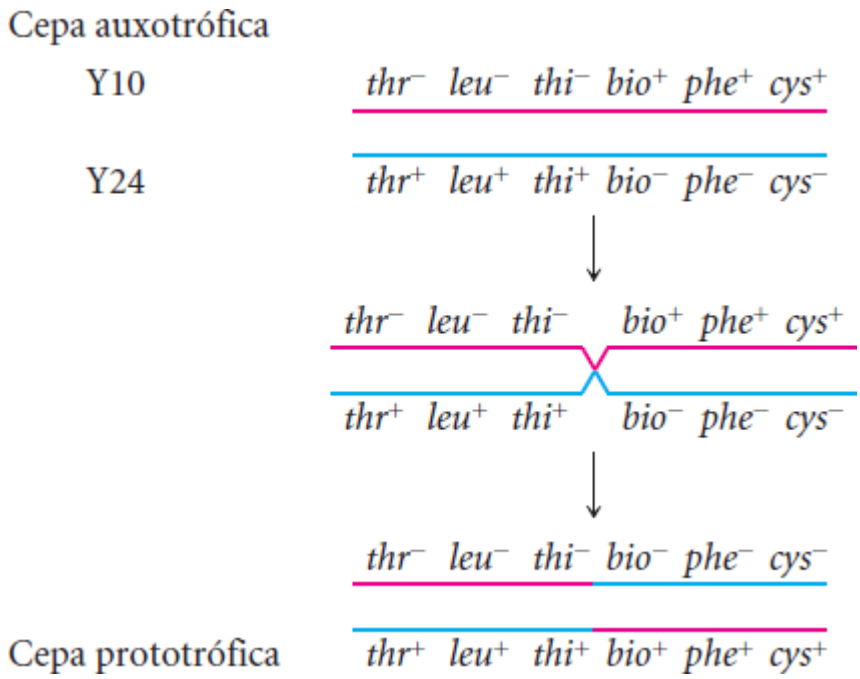


Resultados



Conclusão: Sim, ocorreram troca e recombinação genéticas entre as duas cepas mutantes.

Figura 9.7 O experimento de Lederberg e Tatum demonstrou que as bactérias sofrem troca genética.

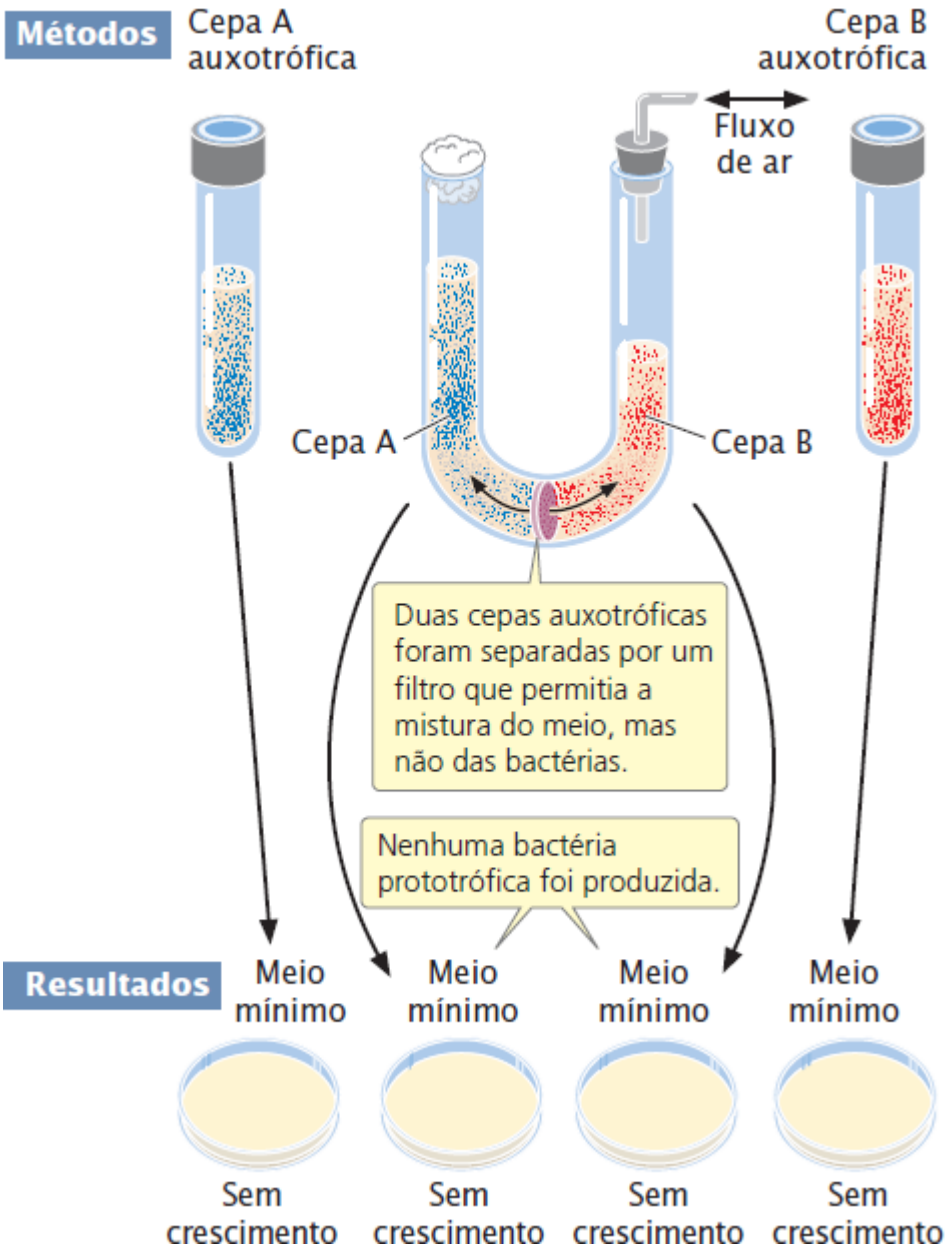


O que eles não sabiam era *como* esse fenômeno tinha ocorrido.

Para estudar esse problema, Bernard Davis criou um tubo em forma de U (**Figura 9.8**) que foi dividido em dois compartimentos por meio de um filtro com poros finos. Esse filtro possibilita que o meio líquido passe de um lado do tubo para o outro, mas os poros do filtro são muito pequenos para permitir a passagem das bactérias. Duas cepas auxotróficas de bactérias foram colocadas em lados opostos do filtro, e aplicou-se sucção alternadamente nas extremidades do tubo em U, fazendo com que o meio fluísse para frente e para trás entre os dois compartimentos. Apesar de horas de incubação no tubo U, as bactérias colocadas em meio mínimo não cresceram; não houve troca genética entre as cepas. Está claro que a troca de genes bacterianos requer o contato direto, ou conjugação, entre as células bacterianas.

Experimento

Pergunta: Como a troca genética observada no experimento de Lederberg e Tatum ocorreu?



Conclusão: A troca genética requer o contato direto entre as células bacterianas.

Figura 9.8 Experimento do tubo em U de Davis.

Células F^+ e F^-. Na maioria das bactérias, a conjugação depende de um fator de fertilidade (F) na célula doadora e ausente na célula receptora. As células com F são chamadas de F^+ e as células sem F, são F^-.

O fator F tem uma origem de replicação e vários genes necessários para conjugação (ver Figura 9.5). Por exemplo, alguns desses genes codificam **pili** sexuais (**pilus**, no singular), que são extensões delgadas da membrana celular. Uma célula com F produz os *pili* sexuais, um deles faz contato com um receptor em uma célula F^- e aproxima as duas células. O DNA é transferido da célula F^+ para a célula F^-. A conjugação pode ocorrer apenas entre uma célula que tem F e uma célula sem F.

Na maioria dos casos, os únicos genes transferidos durante a conjugação entre uma célula F^+ e uma célula F^- são os relacionados com o fator F (**Figura 9.9 A** e **B**). A transferência é iniciada quando uma das fitas de DNA no fator F é cortada em uma origem (*oriT*). Uma extremidade do DNA cortada se separa do círculo e passa para a célula receptora (**Figura 9.9 C**). A replicação ocorre na fita cortada, prosseguindo ao redor do plasmídio circular na célula F^+ e substituindo a fita transferida (**Figura 9.9 D**). Como o plasmídio na célula F^+ é sempre cortado no local *oriT*, esse local sempre entra na célula receptora primeiro, seguido pelo resto do plasmídio. Assim, a transferência de material genético tem uma direção definida. Dentro da célula receptora, a fita única se replica, produzindo uma cópia circular de fita dupla do plasmídio F (**Figura 9.9 E**). Se o fator F inteiro for transferido para a célula F^- receptora, esta se torna uma célula F^+.

Células Hfr. A conjugação transfere material genético no plasmídio F das células F^+ para F^-, mas não é a responsável pela transferência dos genes cromossômicos observados por Lederberg e Tatum. Nas cepas Hfr (alta frequência), o fator F está

integrado ao cromossomo bacteriano (**Figura 9.10**). As células Hfr se comportam como células F^+, formando pelos sexuais e sofrendo conjugação com as células F^-.

Na conjugação entre as células Hfr e F^- (**Figura 9.11 A**), o fator F integrado é cortado e a extremidade da fita cortada se move na direção da célula F^- (**Figura 9.11 B**), como ocorre na conjugação entre as células F^+ e F^-. Como em uma célula Hfr o fator F está integrado ao cromossomo bacteriano, o cromossomo segue na direção da célula receptora. Quanto de cromossomo bacteriano é transferido depende do tempo que as duas células permanecem em conjugação.

Dentro da célula receptora, a fita de DNA doada se replica (**Figura 9.11 C**), e ocorre *crossing over* entre ela e o cromossomo original da célula F^- (**Figura 9.11 D**). Essa transferência de gene entre as células Hfr e F^- é como as células prototróficas recombinantes observadas por Lederberg e Tatum eram produzidas. Após ocorrer *crossing over* na célula receptora, o cromossomo doado é degradado e o cromossomo recombinante do receptor permanece (**Figura 9.11 E**), para ser replicado e transmitido para as próximas gerações por meio de fissão binária.

Em um cruzamento entre Hfr × F^-, a célula F^- quase nunca se torna F^+ ou Hfr porque o fator F é cortado ao meio no início da transferência da fita, colocando parte de F no início e parte no final da fita a ser transferida. Para se tornar Hfr ou F^+, a célula receptora deve receber o fator F inteiro, ou seja, o cromossomo bacteriano inteiro deve ser transferido. É raro ocorrer este caso, porque a maioria das células conjugadas se separa antes de o cromossomo inteiro ser transferido.

O plasmídio F nas células F^+ se integra ao cromossomo bacteriano, fazendo com que a célula F^+ se torne Hfr, em uma frequência de apenas 1 em 10.000. Essa baixa frequência é responsável pela baixa taxa de recombinação observada por Lederberg e Tatum em suas células F^+. O fator F é retirado do cromossomo bacteriano em uma taxa baixa, fazendo com que poucas células Hfr se tornem F^+.

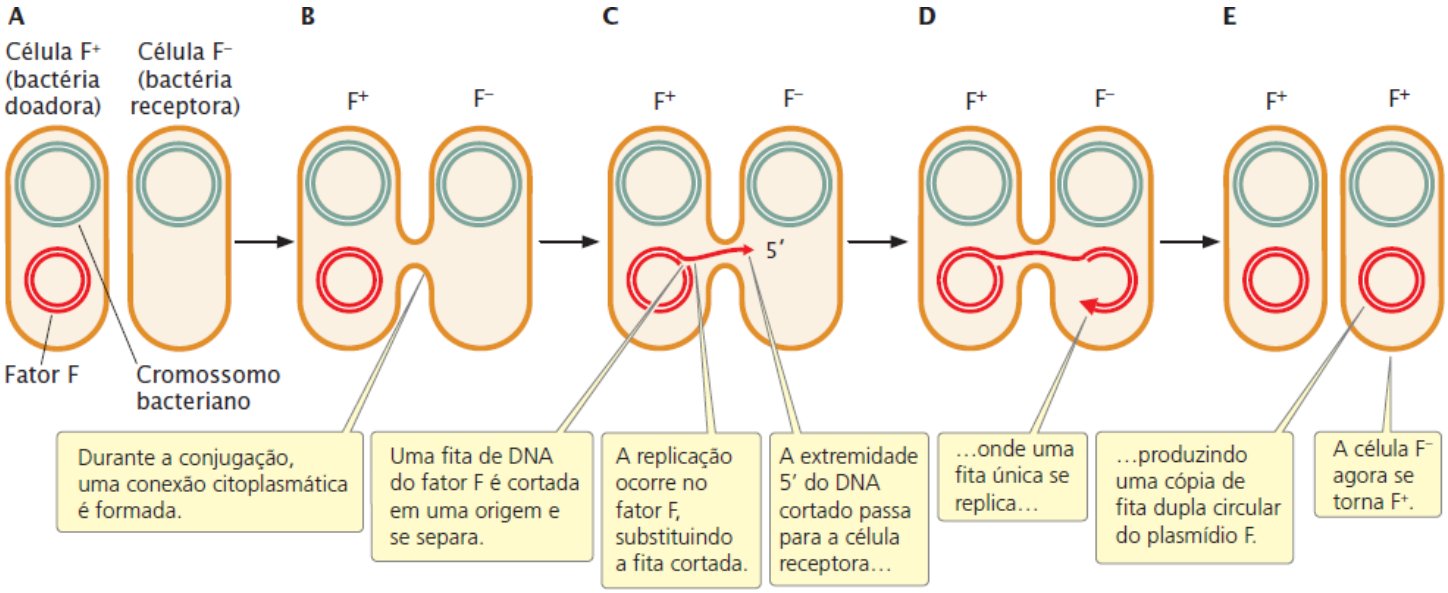


Figura 9.9 O fator F é transferido durante a conjugação entre células F^+ e F^-.

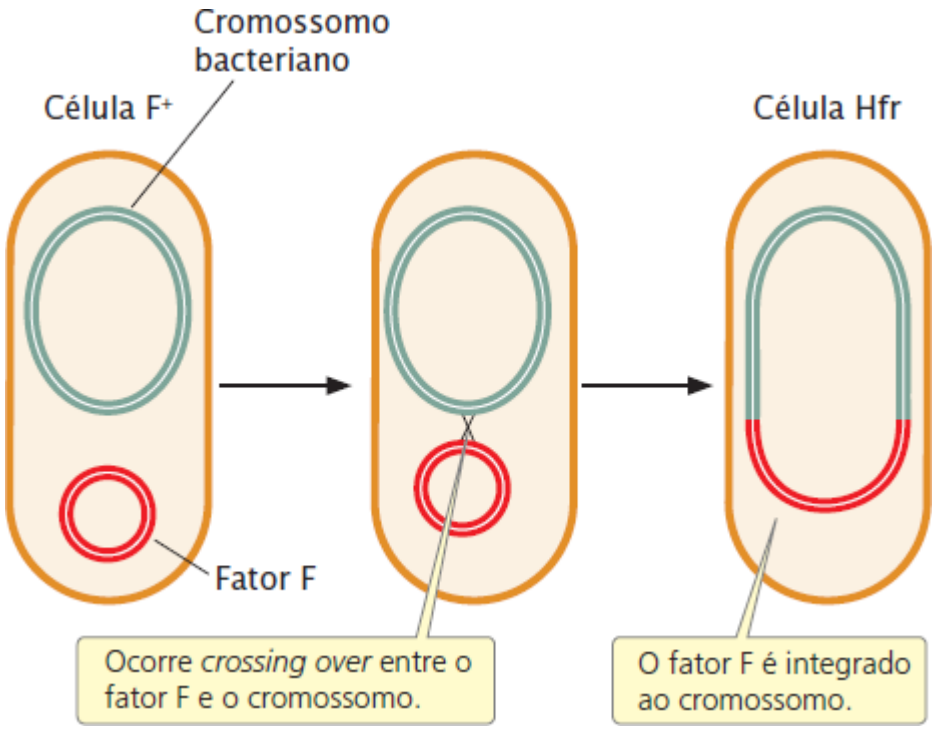


Figura 9.10 **O fator F é integrado ao cromossomo bacteriano em uma célula Hfr.**

Células F'. Quando o fator F é cortado do cromossomo bacteriano, uma pequena quantidade desse cromossomo pode ser removida com ele, e esses genes cromossômicos serão carreados com o plasmídio F (**Figura 9.12**). As células contendo um plasmídio F com alguns genes bacterianos são chamadas de F linha (F′). Por exemplo, se um fator F se integra a um cromossomo adjacente aos genes *lac* (genes que permitem que a célula metabolize o açúcar lactose), o fator F pode pegar os genes *lac* quando for retirado, tornando-se F′ *lac*. As células F′ podem se conjugar com as células F^- porque têm o plasmídio F com todas as informações genéticas necessárias para a conjugação e transferência genética. As características dos diferentes tipos de cruzamento da *E. coli* (células com diferentes tipos de F) estão resumidas no **Quadro 9.2**.

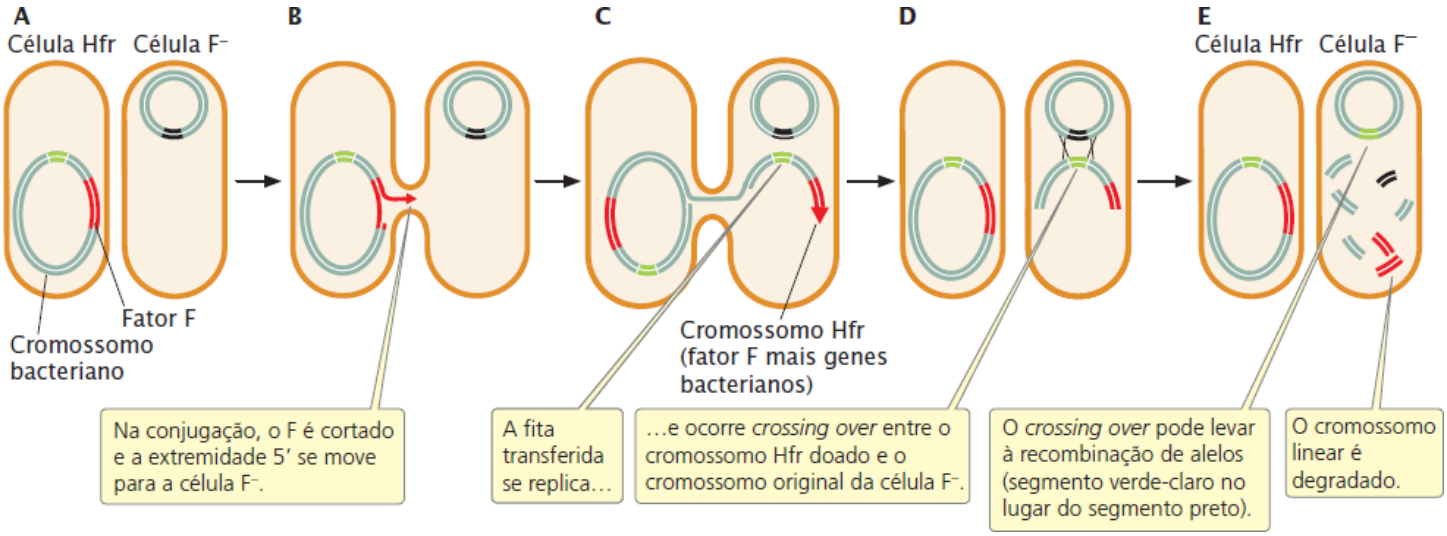


Figura 9.11 **Os genes bacterianos podem ser transferidos de uma célula Hfr para uma célula F^- na conjugação**. Em uma célula Hfr, o fator F está integrado ao cromossomo bacteriano.

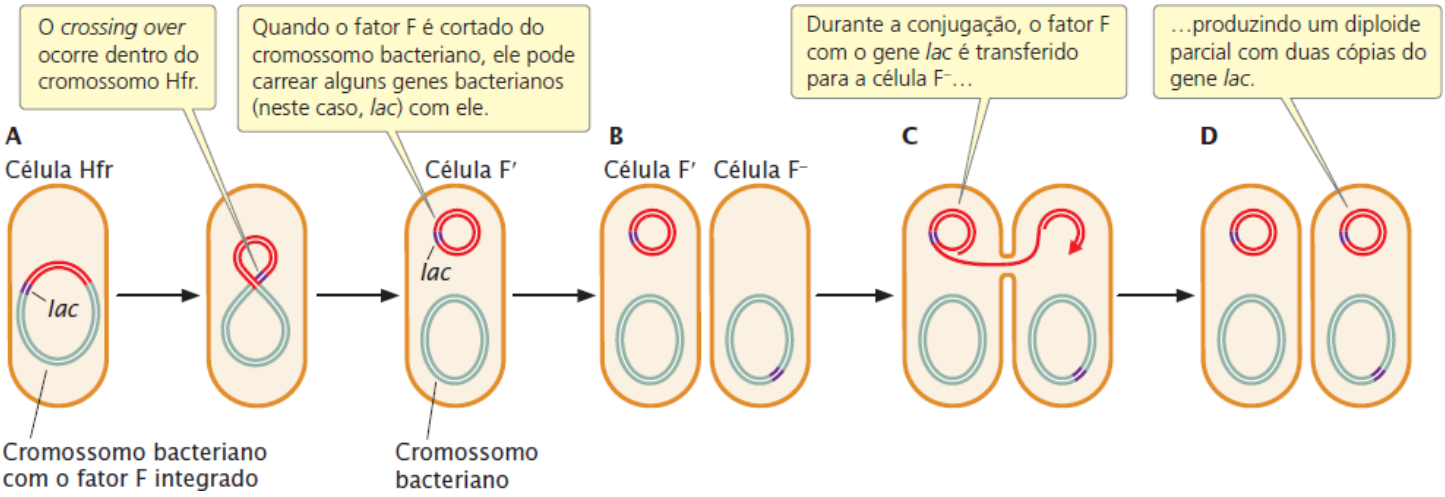


Figura 9.12 Uma célula Hfr pode ser convertida em uma célula F′ quando o fator F é cortado do cromossomo bacteriano e carreia genes bacterianos com ele. A conjugação produz um diploide parcial.

Quadro 9.2 Características das células *E. coli* com diferentes tipos de fator F.

Tipo	Características do fator F	Papel na conjugação
F^+	Presente como DNA circular separado	Doadora
F^-	Ausente	Receptora
Hfr	Presente, integrado ao cromossomo bacteriano	Doadora de alta frequência
F'	Presente como DNA circular separado, carreando alguns genes bacterianos	Doadora

Durante a conjugação entre uma célula F′ *lac* e uma célula F^-, o plasmídio F é transferido para a célula F^-, o que significa que qualquer gene do plasmídio F, incluindo os do cromossomo bacteriano, pode ser transferido para as células F^- receptoras. Esse processo é chamado de sexodução. Ele produz diploides parciais, ou merozigotos, que são células com dois tipos de alguns genes, um no cromossomo bacteriano e um no plasmídio F recém-introduzido. Os desfechos da conjugação entre diferentes tipos de cruzamento da *E. coli* estão resumidos no **Quadro 9.3**.

Conceitos

A conjugação em *E. coli* é controlada por um episssomo chamado fator F. As células contendo F (células F^+) são doadoras durante a transferência genética; as células sem F (células F^-) são receptoras. As células Hfr têm F integrado ao cromossomo bacteriano; elas doam DNA para as células F^- em uma frequência elevada. As células F' têm uma cópia de F com alguns genes bacterianos.

Checagem dos conceitos 3

A conjugação entre uma célula F^+ e F^- em geral resulta em

a. duas células F^+.
b. duas células F^-.
c. uma célula F^+ e uma célula F^-.
d. uma célula Hfr e uma célula F^+.

Como mapear os genes bacterianos com conjugação interrompida. A transferência de DNA que ocorre durante a conjugação entre as células Hfr e F^- permite que os genes bacterianos sejam mapeados. Na conjugação, o cromossomo da célula Hfr é transferido para a célula F^-. A transferência do cromossomo inteiro da *E. coli* leva cerca de 100 min; se a conjugação for interrompida antes de esse período ter transcorrido, apenas parte do cromossomo será transferida para a célula F^- e terá uma oportunidade para recombinar com o cromossomo receptor.

Quadro 9.3 Resultados da conjugação entre células com diferentes fatores F.

Conjugação	Tipos de células conjugadas após a conjugação
$F^+ \times F^-$	Duas células F^+ (a célula F^- se torna F^+)
Hfr × F^-	Uma célula Hfr e uma célula F^- (sem mudança) *
$F' \times F^-$	Duas células F′ (a célula F^- se torna F′)

*Em raros casos, a célula F^- se torna F^+ em uma conjugação Hfr × F^- se o cromossomo inteiro for transferido durante a conjugação.

A transferência de cromossomo sempre inicia dentro do fator F integrado e continua em uma direção contínua, de forma que os genes são transferidos de acordo com sua sequência no cromossomo. O tempo necessário para que os genes sejam transferidos indica suas posições relativas no cromossomo. Na maioria dos mapas genéticos, as distâncias são expressas como porcentagem de recombinação; entretanto, nos mapas bacterianos construídos com a conjugação interrompida, a unidade básica de distância é um minuto.

Problema desenvolvido

Para ilustrar o método de mapeamento de genes com conjugação interrompida, vamos observar um cruzamento analisado por François Jacob e Elie Wollman, que desenvolveram esse método de mapeamento genético (**Figura 9.13 A**). Eles usaram células Hfr doadoras sensíveis ao antibiótico estreptomicina (genótipo *str*^{s}), resistentes à azida sódica (*azi*^{r}) e à infecção por bacteriófago T1 (*ton*^{r}), prototróficas para treonina (*thr*$^+$) e leucina (*leu*$^+$) e capazes de degradar lactose (*lac*$^+$) e galactose (*gal*$^+$). Usaram também células F^- receptoras resistentes à estreptomicina (*str*^{r}), sensíveis à azida sódica (*azi*^{s}) e à infecção por bacteriófago T1 (*ton*^{s}), auxotróficas para treonina (*thr*$^-$) e leucina (*leu*$^-$) e incapazes de degradar lactose (*lac*$^-$) e galactose (*gal*$^-$). Assim, os genótipos das células doadoras e receptoras eram:

Células Hfr doadoras: *str*^{s} *leu*$^+$ *thr*$^+$ *azi*^{r} *ton*^{r} *lac*$^+$ *gal*$^+$

Células F^- receptoras: *str*^{r} *leu*$^-$ *thr*$^-$ *azi*^{s} *ton*^{s} *lac*$^-$ *gal*$^-$

As duas cepas foram misturadas em um meio nutriente e deixadas para conjugar. Após alguns minutos, o meio era diluído para evitar novos pareamentos. Em intervalos regulares, uma amostra de células era removida e agitada vigorosamente em um liquidificador para interromper qualquer conjugação e transferência de DNA. As células de cada amostra eram plaqueadas em um meio seletivo com estreptomicina e sem leucina e treonina. As células doadoras originais eram sensíveis à estreptomicina (*str*^{s}) e não cresceriam nesse meio. As células F^- receptoras eram auxotróficas para leucina e treonina e também não cresceriam nesse meio. Apenas células que sofriam conjugação e recebiam pelo menos os genes *leu*$^+$ e *thr*$^+$ das doadoras Hfr poderiam crescer nesse meio. Todas as células *str*^{r} *leu*$^+$ *thr*$^+$ foram testadas quanto à existência de outros genes que poderiam ter sido transferidos da cepa Hfr doadora.

Como Jacob e Wollman usaram estreptomicina para exterminar todas as células doadoras, eles não puderam examinar a transferência do gene *str*^{s}. Todas as células que cresceram no meio seletivo eram *leu*$^+$ *thr*$^+$, então sabemos que esses genes foram transferidos. Foi traçado um gráfico da porcentagem de células *str*^{r} *leu*$^+$ *thr*$^+$ que receberam alelos específicos (*azi*^{r}, *ton*^{r}, *leu*$^+$ e *gal*$^+$) da cepa Hfr e a duração da conjugação (**Figura 9.13 B**). Em que ordem os genes foram transferidos e quais são as distâncias entre eles?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

Da ordem dos genes no cromossomo bacteriano e a distância entre eles.

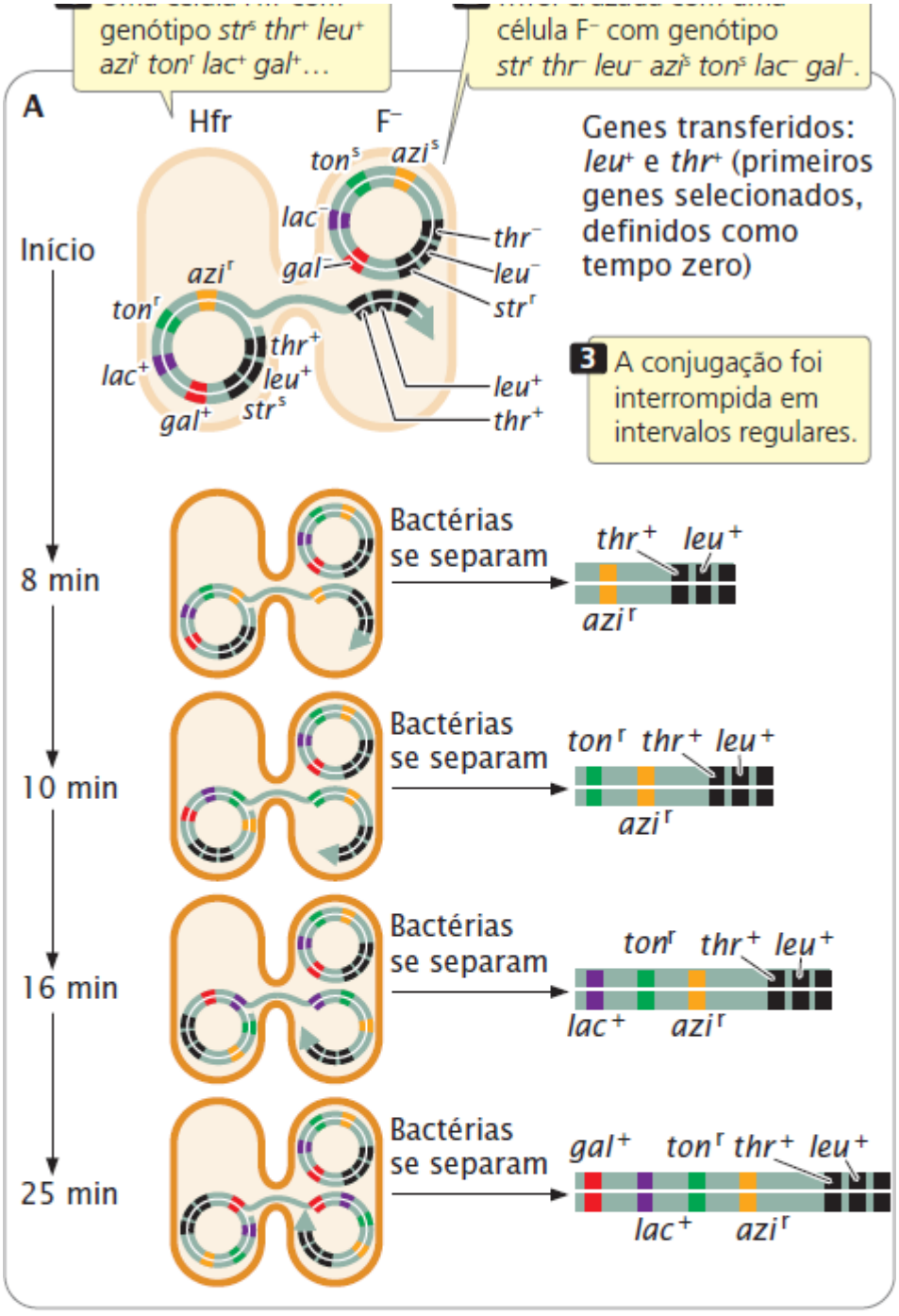
Experimento

Pergunta: Como a conjugação interrompida pode ser usada para mapear os genes bacterianos?

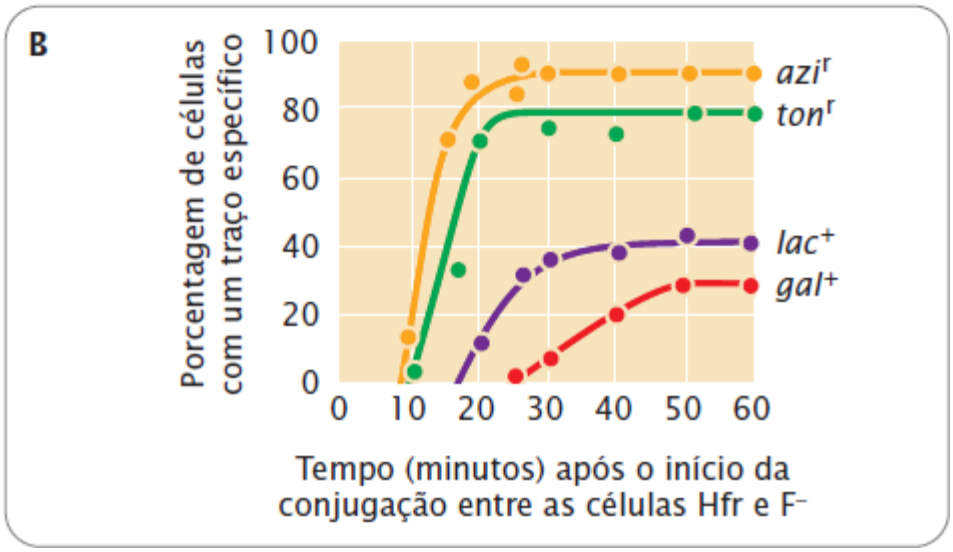
Métodos

1 Uma célula Hfr com

2 ...foi cruzada com uma


genótipo str^s thr^+ leu^+ azi^s ton^r lac^+ gal^+...
célula F^– com genótipo str^r thr^– leu^– azi^s ton^s lac^– gal^–.
A
Hfr
F^–
Genes transferidos: leu^+ e thr^+ (primeiros genes selecionados, definidos como tempo zero)
Início
3 A conjugação foi interrompida em intervalos regulares.
8 min
10 min
16 min
25 min
Bactérias se separam
thr^+ leu^+ azi^r
ton^r thr^+ leu^+ azi^r
ton^r thr^+ leu^+ lac^+ azi^r
gal^+ ton^r thr^+ leu^+ lac^+ azi^r

Resultados


B
Porcentagem de células com um traço específico
100
80
60
40
20
0
0 10 20 30 40 50 60
azi^r
ton^r
lac^+
gal^+
Tempo (minutos) após o início da conjugação entre as células Hfr e F^–

Conclusão: Os tempos de transferência indicam a ordem e as distâncias relativas entre os genes e podem ser usados para construir um mapa genético.

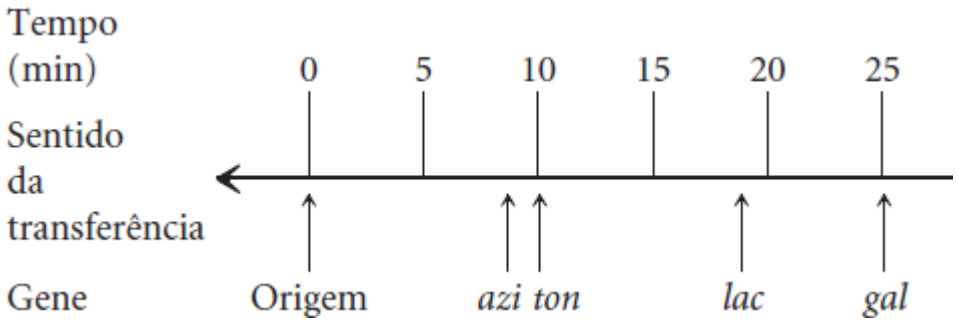
Figura 9.13 Jacob e Wollman usaram a conjugação interrompida para mapear genes das bactérias.

Quais informações são fornecidas para solucionar o problema?

- As células doadoras eram *str*^{s} *leu*$^{+}$ *thr*$^{+}$ *azi*^{r} *ton*^{r} *lac*$^{+}$ *gal*$^{+}$, e as células receptoras eram *str*^{r} *leu*$^{-}$ *thr*$^{-}$ *azi*^{s} *ton*^{s} *lac*$^{-}$ *gal*$^{-}$
- A porcentagem de células receptoras com diferentes traços que aparecem em tempos diferentes após o início da conjugação (ver Figura 9.13).

Etapas para a solução

O primeiro gene doador a aparecer em todas as células receptoras (em cerca de 9 min) era *azi*^{r}. O gene *ton*^{r} aparece logo depois (cerca de 10 min), seguido por *lac*$^{+}$ (cerca de 18 min) e *gal*$^{+}$ (depois de 25 min). Esses tempos de transferência indicam a ordem da transferência dos genes e as distâncias relativas entre eles (ver Figura 9.13 B).



Observe que a frequência da transferência de genes das células doadoras para as receptoras caiu nos genes mais distantes. Por exemplo, cerca de 90% das células receptoras receberam o alelo *azi*^{r}, mas apenas 30% receberam o alelo *gal*$^{+}$. A porcentagem menor para *gal*$^{+}$ ocorre porque algumas células conjugadas se separaram espontaneamente antes de serem rompidas pelo liquidificador. A probabilidade de ruptura espontânea aumenta com o tempo, então menos células tiveram a oportunidade de receber os genes que foram transferidos mais tarde.

Para treinar mais sobre o mapeamento de genes bacterianos com a conjugação interrompida, resolva o **Problema 23** no final do capítulo.

Transferência direcionada e mapeamento. Diferentes cepas Hfr de uma dada espécie de bactéria têm o fator F integrado ao cromossomo bacteriano em diferentes locais e em diferentes orientações. A transferência genética sempre começa dentro do F, e a orientação e a posição do F determinam a direção e o ponto de início da transferência dos genes. A **Figura 9.14 A** mostra que, na cepa Hfr1, F está integrado entre *leu* e *azi*; a orientação de F nesse local dita que a transferência de genes será feita no sentido anti-horário ao redor do cromossomo circular. Os genes dessa cepa serão transferidos na ordem:

←*leu-thr-thi-his-gal-lac-pro-azi*

Na cepa Hfr5, F está integrado entre os genes *thi* e *his* (**Figura 9.14 B**) e na orientação oposta. Aqui, a transferência de genes será no sentido horário:

←*thi-thr-leu-azi-pro-lac-gal-his*

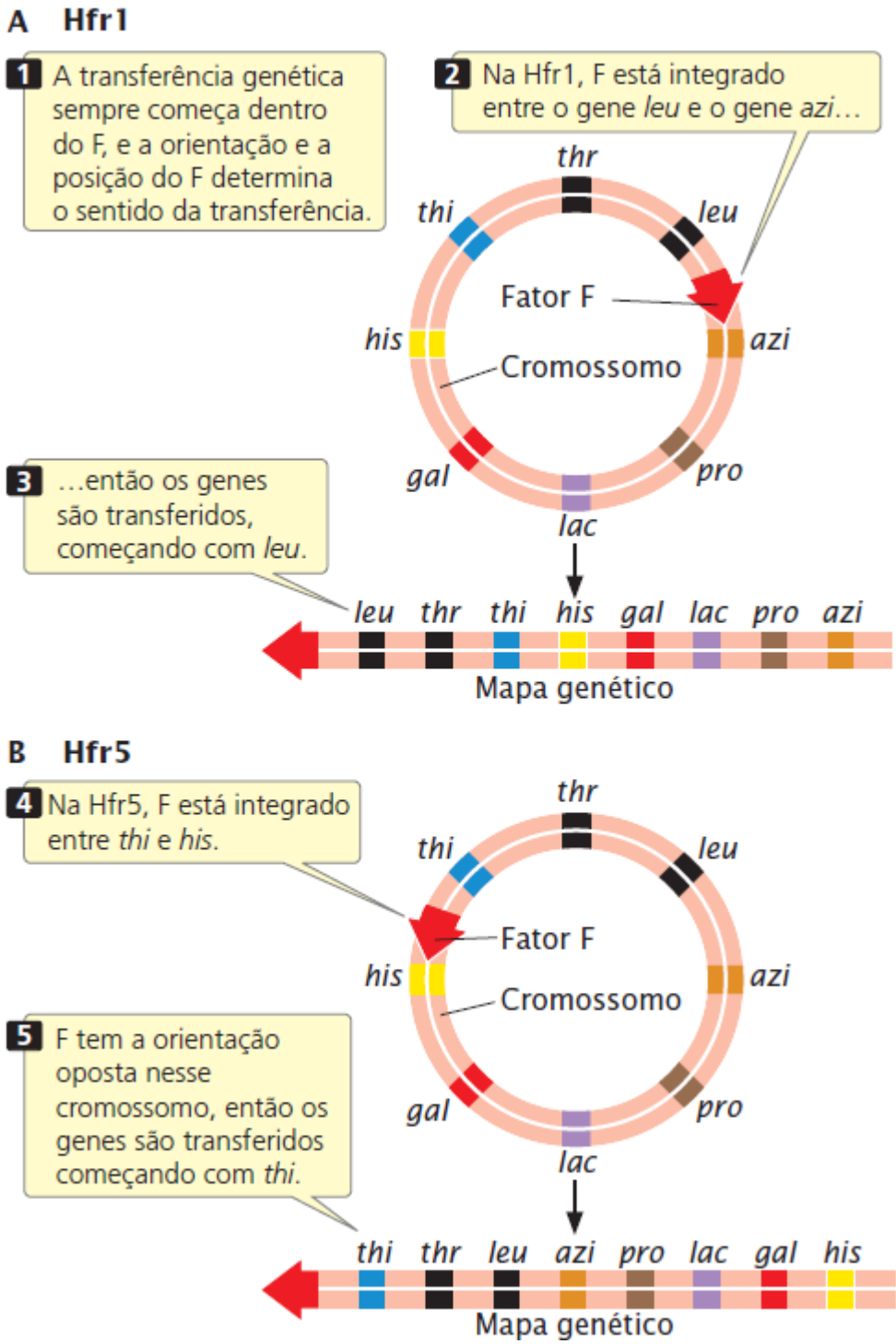


Figura 9.14 A orientação do fator F em uma cepa Hfr determina o sentido da transferência genética. As pontas de setas indicam a origem e o sentido da transferência.

Embora o ponto de início e a direção da transferência possam ser diferentes entre duas cepas, a distância relativa no momento entre dois pares de genes é constante.

Observe que a ordem da transferência de genes não é a mesma para diferentes cepas Hfr (**Figura 9.15 A**). Por exemplo, o gene *azi* é transferido logo após o gene *leu* na cepa HfrH, mas muito depois de *leu* na cepa HFr1. O alinhamento das sequências (**Figura 9.15 B**) mostra que os dois genes ao lado de *azi* são sempre os mesmos: *leu* e *pro*. Que eles são o mesmo faz sentido quando identificamos que o cromossomo bacteriano é circular e o ponto de início da transferência varia entre as cepas. Esses dados fornecem a primeira evidência de que o cromossomo bacteriano é circular (**Figura 9.15 C**). Resolva o Problema 22

Conceitos

A conjugação pode ser usada para mapear genes bacterianos ao misturar células Hfr e F^- de diferentes genótipos e interromper a conjugação em intervalos regulares. O tempo necessário para os genes serem transferidos das células Hfr para F^- indica as posições relativas dos genes no cromossomo bacteriano.

Checagem dos conceitos 4

A conjugação interrompida foi usada para mapear três genes na *E. coli.* Os genes doadores aparecem primeiro nas células receptoras nos seguintes tempos: *gal*, 10 minutos; *his*, 8 minutos e *pro*, 15 minutos. Qual gene está no meio?

Transferência natural dos genes e resistência a antibióticos

Os antibióticos são substâncias usadas para exterminar as bactérias. Seu desenvolvimento e seu uso disseminado reduziram de forma significativa a ameaça de infecções e salvaram inúmeras vidas. No entanto, muitas bactérias patogênicas desenvolveram resistência aos antibióticos, particularmente em ambientes onde os antibióticos são usados na rotina, como hospitais, criações de animais e fazenda de criação de peixes. Nesses ambientes, as únicas bactérias que sobrevivem são as que têm resistência aos antibióticos. Sem a competição com outras bactérias, rapidamente as resistentes se multiplicam e se espalham. Assim, os antibióticos selecionam as bactérias resistentes e reduzem a eficácia do tratamento antibiótico para infecções.

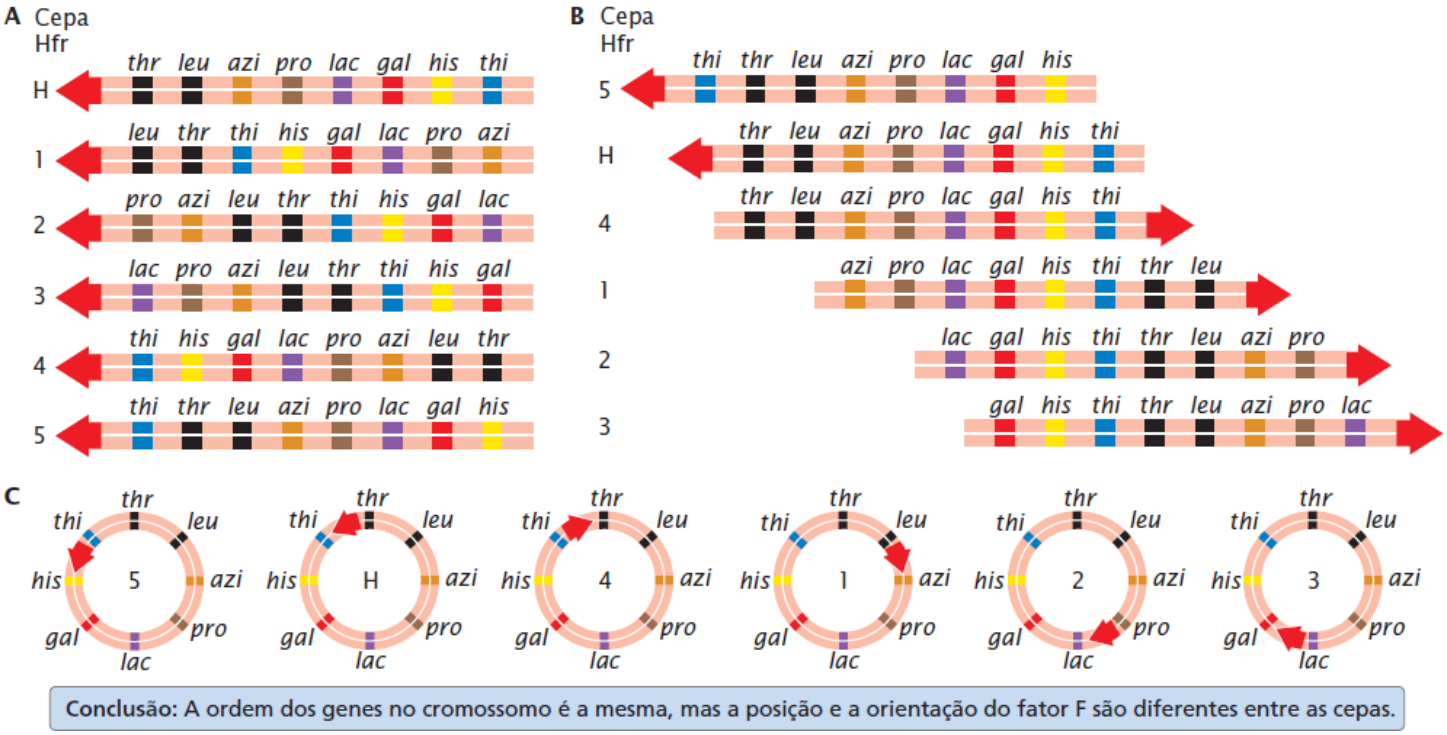


Figura 9.15 A ordem da transferência dos genes em várias cepas Hfr diferentes indica que o cromossomo da *E. coli* é circular.

A resistência a antibióticos nas bactérias resulta, com frequência, da ação de genes localizados nos *plasmídios R*, pequenos plasmídios circulares que podem ser transferidos por conjugação. Os plasmídios R de farmacorresistência evoluíram nos últimos 60 anos (desde o início do uso disseminado dos antibióticos), e alguns deles conduzem resistência a vários antibióticos simultaneamente. Curioso, mas plausível, é que as fontes de alguns dos genes de resistência encontrados nos plasmídios R são os micróbios que produzem antibióticos em primeiro lugar. Os plasmídios R podem se espalhar com facilidade pelo ambiente, passando entre bactérias próximas ou não em várias situações comuns. Por exemplo, a pesquisa mostra que os plasmídios que carreiam genes para a resistência a múltiplos antibióticos foram transferidos de uma vaca infectada com *E. coli* para uma cepa humana de *E. coli* em uma toalha de mão: um fazendeiro que seca suas mãos após ordenhar uma vaca infectada pode inconscientemente transferir a resistência a antibiótico de micróbios bovinos para micróbios existentes nos seres humanos. A conjugação que ocorre na carne picada em uma tábua permite que os plasmídios R sejam transferidos da *E. coli* suína para *E. coli* humana. A transferência dos plasmídios R também ocorre no esgoto, na água dos lagos e em sedimentos marinhos. O fato de os plasmídios R poderem ser disseminados facilmente pelo ambiente e transmitidos entre bactérias próximas ou não enfatiza a importância de limitar o uso de antibióticos para o tratamento das infecções e a importância da higiene no dia a dia.

Transformação em bactérias

A segunda forma na qual o DNA pode ser transferido entre bactérias é pela transformação (ver Figura 9.6 B). A transformação foi importante na identificação inicial do DNA como material genético, o que será discutido no Capítulo 10.

A transformação requer a captação de DNA a partir do meio circundante e sua incorporação no cromossomo bacteriano ou

em um plasmídio. Ela pode ocorrer de forma natural, quando a bactéria morta se rompe e libera seu DNA para o ambiente. No solo e no ambiente marinho, essa forma pode ser uma via importante de troca de material genético para algumas bactérias.

Mecanismo de transformação. As células que captam DNA por meio de suas membranas são chamadas de **células competentes**. Algumas espécies de bactérias captam DNA mais facilmente do que outras: a competência é influenciada pelo estágio de crescimento, pela concentração de DNA disponível no ambiente e por outros fatores ambientais. O DNA que a célula competente capta não precisa ser bacteriano: praticamente qualquer tipo de DNA (bacteriano ou outro) pode ser captado por células competentes sob condições adequadas.

Quando um fragmento de DNA penetra na célula durante a transformação (**Figura 9.16**), uma das fitas é degradada, enquanto a outra fita atravessa a membrana e pode parear com uma região homóloga e ser integrada ao cromossomo bacteriano. Esta integração requer dois eventos de *crossing over*, após os quais o DNA de fita única remanescente é degradado pelas enzimas bacterianas. Em algumas bactérias, o DNA de fita dupla atravessa a membrana celular e é integrado ao cromossomo bacteriano.

Os geneticistas que estudam as bactérias desenvolveram técnicas para aumentar a frequência de transformação no laboratório para produzir fragmentos específicos de DNA nas células. Eles também desenvolveram cepas de bactérias que são mais competentes do que as células selvagens. O tratamento com cloreto de cálcio, choque com calor ou campo elétrico torna as membranas bacterianas mais porosas e permeáveis ao DNA, e a eficiência da transformação também pode ser aumentada ao usar altas concentrações de DNA. Essas técnicas permitem que os pesquisadores transformem bactérias como *E. coli*, que não são naturalmente competentes.

Mapeamento genético com transformação. A transformação, tal como a conjugação, é usada para mapear genes bacterianos, em especial nas espécies que não sofrem conjugação ou transdução (ver Figura 9.6 A e C). O mapeamento por transformação requer duas cepas de bactérias com vários traços genéticos diferentes; por exemplo, a cepa receptora pode ser $a^- b^- c^-$ (auxotrófica para três nutrientes), e a célula doadora prototrófica com os alelos $a^+ b^+ c^+$ (**Figura 9.17**). O DNA da cepa doadora é isolado, purificado e fragmentado. A cepa receptora é tratada para aumentar sua competência, e o DNA da cepa doadora é adicionado ao meio. Os fragmentos do DNA doado entram nas células receptoras e sofrem recombinação com as sequências homólogas de DNA no cromossomo bacteriano. As células que recebem material genético por meio da transformação são chamadas de **transformantes**.

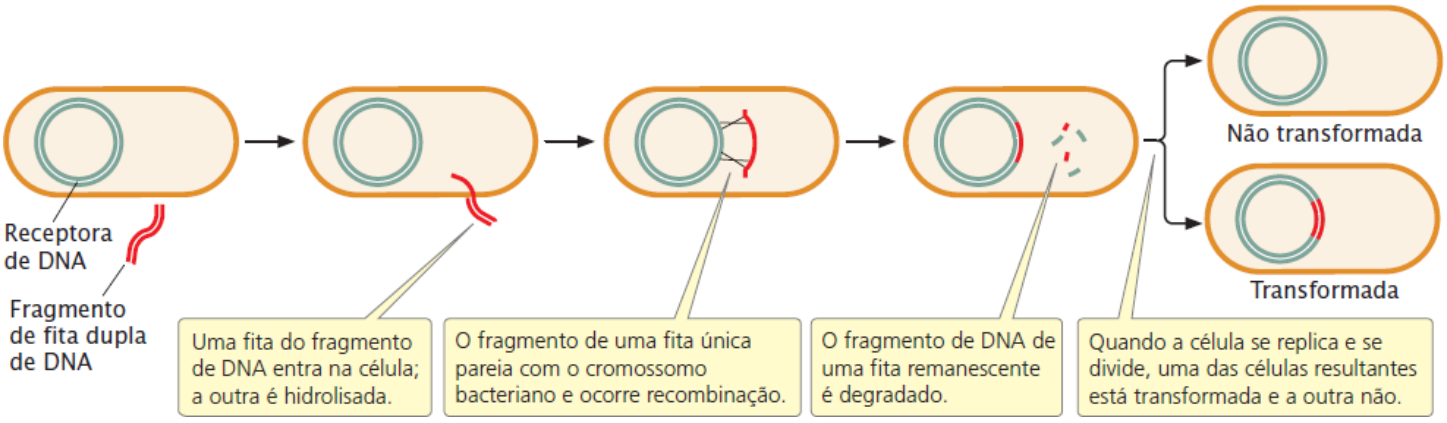


Figura 9.16 **Os genes podem ser transferidos entre bactérias pela transformação.**

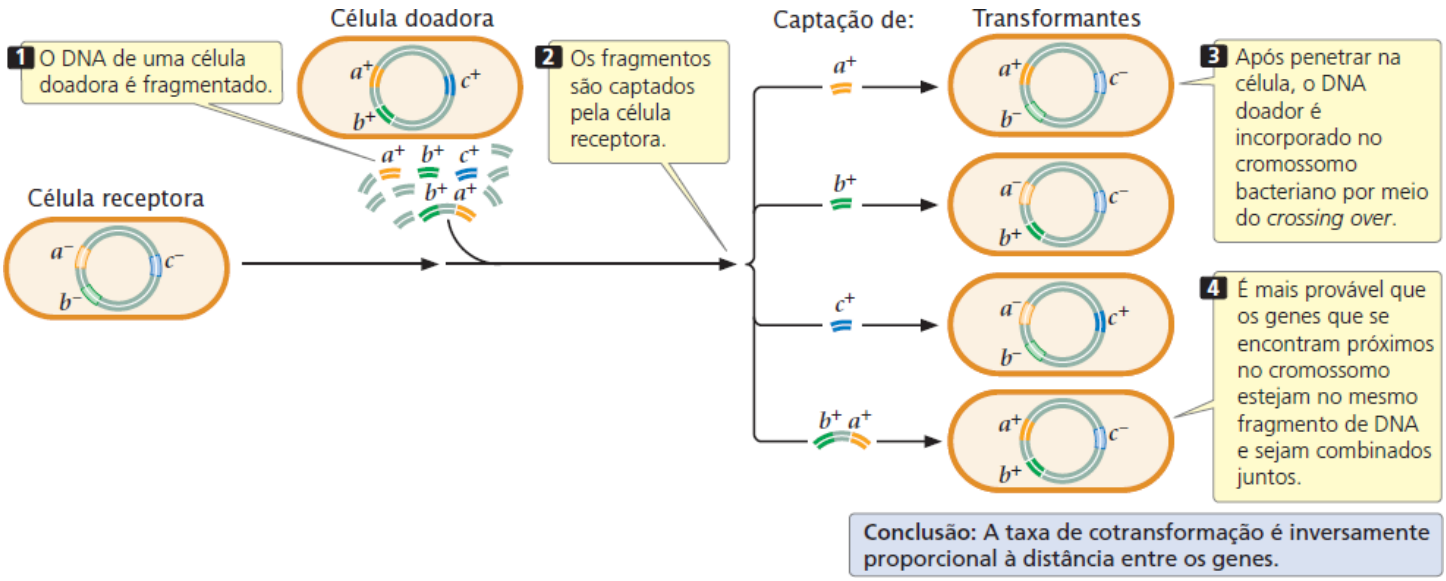


Figura 9.17 A transformação pode ser usada para mapear os genes bacterianos.

Os genes podem ser mapeados ao observar a taxa na qual dois ou mais genes são transferidos para o cromossomo hospedeiro, ou **cotransformados**, na transformação. Quando o DNA doado é fragmentado antes da transformação, é mais provável que os genes que estão fisicamente próximos no cromossomo estejam no mesmo fragmento de DNA e sejam transferidos juntos, como é demonstrado para os genes a^+ e b^+ na Figura 9.17. É improvável que genes distantes estejam no mesmo fragmento de DNA, e em raros casos serão transferidos juntos. Dentro da célula, o DNA é incorporado ao cromossomo bacteriano por meio da recombinação. Se dois genes estiverem muito próximos no mesmo fragmento, é provável que ocorram dois *crossing overs* em cada lado dos dois genes, possibilitando que ambos se tornem parte do cromossomo receptor. Se dois genes estiverem distantes, pode ocorrer um *crossing over* entre eles, permitindo que um gene, mas não o outro, recombine com o cromossomo bacteriano. Assim, é mais provável que dois genes sejam transferidos juntos quando estão próximos no cromossomo, e é raro os genes distantes serem cotransformados. Portanto, a frequência da cotransformação pode ser usada para mapear os genes bacterianos. Se os genes *a* e *b*, assim como os genes *b* e *c*, forem cotransformados com frequência, embora seja raro os genes *a* e *c* serem cotransformados, então o gene *b* está obrigatoriamente entre *a* e *c* – a ordem do gene é *a b c*. Resolva o Problema 25

Conceitos

Os genes podem ser mapeados nas bactérias usando a transformação – a capacidade das células em captar DNA do ambiente e incorporá-lo a seus cromossomos por meio do *crossing over*. A taxa relativa na qual os pares de genes são cotransformados indica a distância entre eles: quanto maior a taxa de cotransformação, mais próximos os genes estão no cromossomo bacteriano.

Checagem dos conceitos 5

O DNA de uma cepa de bactéria com genótipo *his*$^-$ *leu*$^-$ *thr*$^-$ é transformado com o DNA de uma cepa que é *his*$^+$ *leu*$^+$ *thr*$^+$. São encontradas algumas células *leu*$^+$ *thr*$^+$ e outras *his*$^+$ *thr*$^+$, mas não foram observadas células *his*$^+$ *leu*$^+$. Quais genes estão mais distantes?

Sequências de genoma bacteriano

Os mapas genéticos servem como base para informações mais detalhadas fornecidas pelo sequenciamento de DNA, como o conteúdo e a organização dos genes (ver *Capítulo 19* para uma discussão sobre sequenciamento de genes). Agora os geneticistas determinaram as sequências completas de nucleotídios de mais de dois mil genomas bacterianos (ver Quadro 20.1), e vários projetos de sequenciamento de genomas microbianos estão em andamento.

Características dos genomas bacterianos. A maioria dos genomas bacterianos tem de 1 milhão a 4 milhões de pares de bases de DNA, mas alguns são muito menores (p. ex., 580.000 pb no *Mycoplasma genitalium*) e alguns são bem maiores (p. ex., mais de 7 milhões de pb no *Mesorhizobium loti*). Acredita-se que esse tamanho pequeno de genomas bacterianos relativos aos encontrados nos eucariotos multicelulares, que em geral têm bilhões de pares de bases de DNA, seja uma adaptação para a rápida divisão celular, porque a taxa de divisão celular é limitada pelo tempo necessário para replicar o DNA. Por outro lado, a falta de mobilidade na maioria das bactérias demanda flexibilidade metabólica e ambiental, e é provável que o tamanho e o

conteúdo do genoma sejam um equilíbrio entre as forças evolutivas opostas da perda do gene para manter a rápida reprodução e aquisição de genes para manter a flexibilidade.

A função de uma quantidade substancial de genes em todas as bactérias ainda não foi determinada. Alguns genes, em particular aqueles com funções relacionadas, tendem a estar localizados próximos uns dos outros, mas esses grupos estão em locais muitos diferentes em diferentes espécies, o que sugere que os genes bacterianos estão sendo constantemente misturados. As comparações das sequências de genes em bactérias patógenas e benignas têm ajudado a identificar genes relacionados à doença, e podem surgir novos alvos para os antibióticos e outros agentes antimicrobianos.

Transferência horizontal de genes

A disponibilidade de sequências de genomas forneceu indícios que muitas bactérias adquiriram informações genéticas de outras espécies de bactérias e, às vezes, até de organismos eucarióticos em um processo chamado de **transferência horizontal de genes**. Na maioria dos eucariotos, os genes são transferidos apenas entre membros da mesma espécie pelo processo de reprodução (chamado de transmissão vertical); ou seja, os genes são transferidos de uma geração para a próxima. Na transferência horizontal, os genes podem ser transferidos entre membros individuais de diferentes espécies por mecanismos não reprodutivos, como conjugação, transformação e transdução. A evidência sugere que a transferência horizontal de genes se repete entre as bactérias. Por exemplo, até 17% do genoma da *E. coli* foram adquiridos de outras bactérias por meio da transferência horizontal dos genes. Algumas bactérias patogênicas de importância médica adquiriram os genes necessários para infecção, enquanto outras adquiriram genes que conferem resistência aos antibióticos.

Graças à ocorrência disseminada de transferência horizontal, muitos cromossomos bacterianos são uma mistura de genes herdados por transmissão vertical e genes adquiridos por meio dessa transferência. Essa situação fez alguns biólogos questionarem se o conceito de espécie é adequado para as bactérias. Uma espécie é definida como um grupo de organismos isolados reprodutivamente de outros grupos, com um conjunto de genes em comum e que se desenvolvem juntos (ver *Capítulo 26*). Graças à transferência horizontal de genes, os genes de uma espécie bacteriana não são isolados a partir dos genes de outra espécie, dificultando aplicar o conceito tradicional de espécie. A transferência horizontal também confunde a determinação das relações ancestrais entre as bactérias. A reconstrução de relações ancestrais baseia-se nas semelhanças e diferenças genéticas: considera-se que os organismos geneticamente semelhantes descendem de um ancestral recente comum, enquanto aqueles geneticamente diferentes têm uma relação mais distante. Entretanto, por meio da transferência horizontal de genes, até bactérias distantes podem ter genes em comum e parece que descendem de um ancestral recente comum. Atualmente, a natureza da espécie e como classificar as bactérias são tópicos controversos dentro do campo da microbiologia.

9.3 Vírus são sistemas simples, replicantes, receptivos à análise genética

Todos os organismos – plantas, animais, fungos e bactérias – são infectados por vírus. Um **vírus** é uma estrutura replicante simples composta de ácido nucleico circundado por um revestimento proteico (ver Figura 2.3). Os vírus têm grande variedade de formatos e tamanhos (**Figura 9.18**). Alguns têm DNA como seu material genético, enquanto outros têm RNA; o ácido nucleico pode ser de fita dupla ou única, linear ou circular.

Os vírus que infectam as bactérias (bacteriófagos ou fagos, para resumir) são cruciais na pesquisa genética desde o final dos anos 1940. Eles são ideais para muitos tipos de pesquisa genética porque são pequenos e têm genomas facilmente manipuláveis, reproduzem-se com rapidez e produzem vários descendentes. Os bacteriófagos têm dois ciclos de vida alternativos: os ciclos *lítico* e *lisogênico*. No ciclo lítico, um fago se prende a um receptor na parede celular da bactéria e injeta seu DNA na célula (**Figura 9.19**). Dentro da célula hospedeira, o DNA do fago é replicado, transcrito e traduzido, produzindo mais DNA e proteínas do fago. Novas partículas do fago são montadas a partir desses componentes. Os fagos produzem uma enzima que abre a célula hospedeira, liberando novos fagos. Os **fagos virulentos** se reproduzem basicamente por meio do ciclo lítico e sempre matam suas células hospedeiras.

Os **fagos temperados** podem sofrer o ciclo lítico ou lisogênico. O ciclo lisogênico inicia como o ciclo lítico (ver Figura 9.19), mas, dentro da célula, o DNA do fago se integra ao cromossomo bacteriano, onde permanece como um **prófago** inativo. O prófago é replicado junto com o DNA bacteriano e transmitido quando a bactéria se divide. Alguns estímulos podem fazer com que o prófago se dissocie do cromossomo bacteriano e entre no ciclo lítico, produzindo novas partículas de fago e rompendo a célula.

Técnicas para o estudo dos bacteriófagos

Os vírus se reproduzem apenas dentro das células hospedeiras, de modo que os bacteriófagos têm de ser cultivados nas células

bacterianas. Os fagos podem ser cultivados em grandes culturas líquidas de bactérias para gerar muitos descendentes, mas é preciso isolá-los em placas se quisermos estudar as características do descendente. Misturamos fagos e bactérias e plaqueamos no meio sólido em uma placa de Petri. É usada uma concentração elevada de bactérias para que as colônias cresçam em outra e produzam uma camada contínua de bactérias, ou "campo" no ágar. Um fago individual infecta uma única célula bacteriana e entra em seu ciclo lítico. Muitos novos fagos são liberados da célula rompida e infectam outras células; o ciclo, então, se repete. Como as bactérias crescem em meio sólido, a difusão dos fagos é restrita e apenas as células próximas são infectadas. Após vários ciclos de reprodução do fago, aparece uma clara mancha de células rompidas, uma **placa**, na placa de Petri (**Figura 9.20**). Cada placa representa um único fago que se multiplicou e rompeu várias células. Plaquear um volume conhecido de uma solução diluída de fagos em uma camada bacteriana e contar o número de placas que surgem pode ser considerado para determinar a concentração original de fagos na solução.

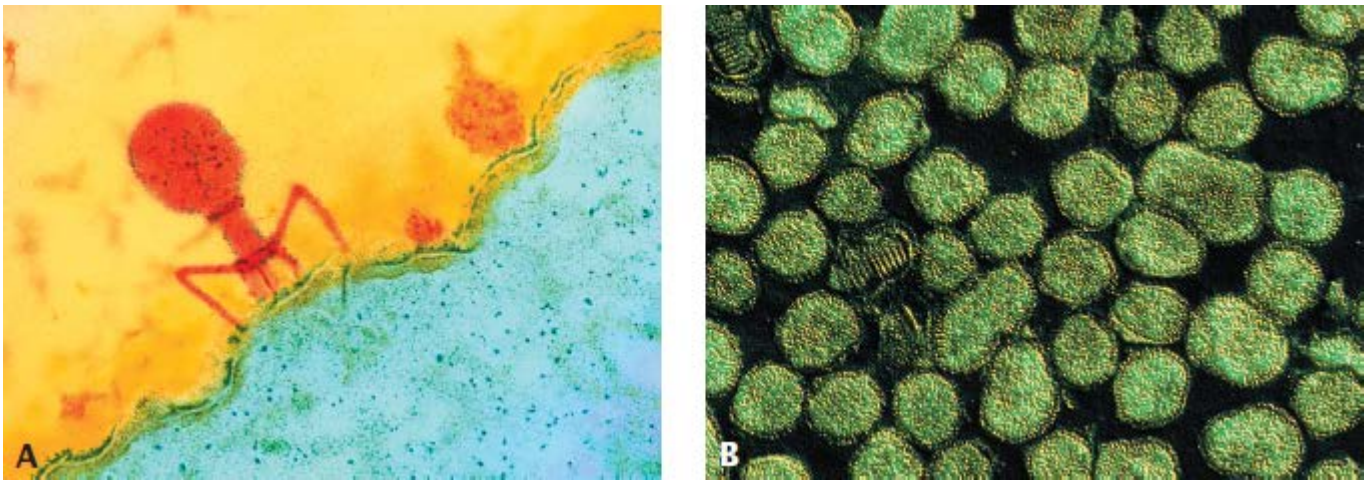


Figura 9.18 Os vírus têm diferentes estruturas e tamanhos. A. Bacteriófago T4 (*laranja brilhante*). **B.** Vírus influenza A (*estruturas em verde*). (*Parte A*: Biozentrum, University of Basel/Photo Researchers. *Parte B*: Eye of Science/Photo Researchers.)

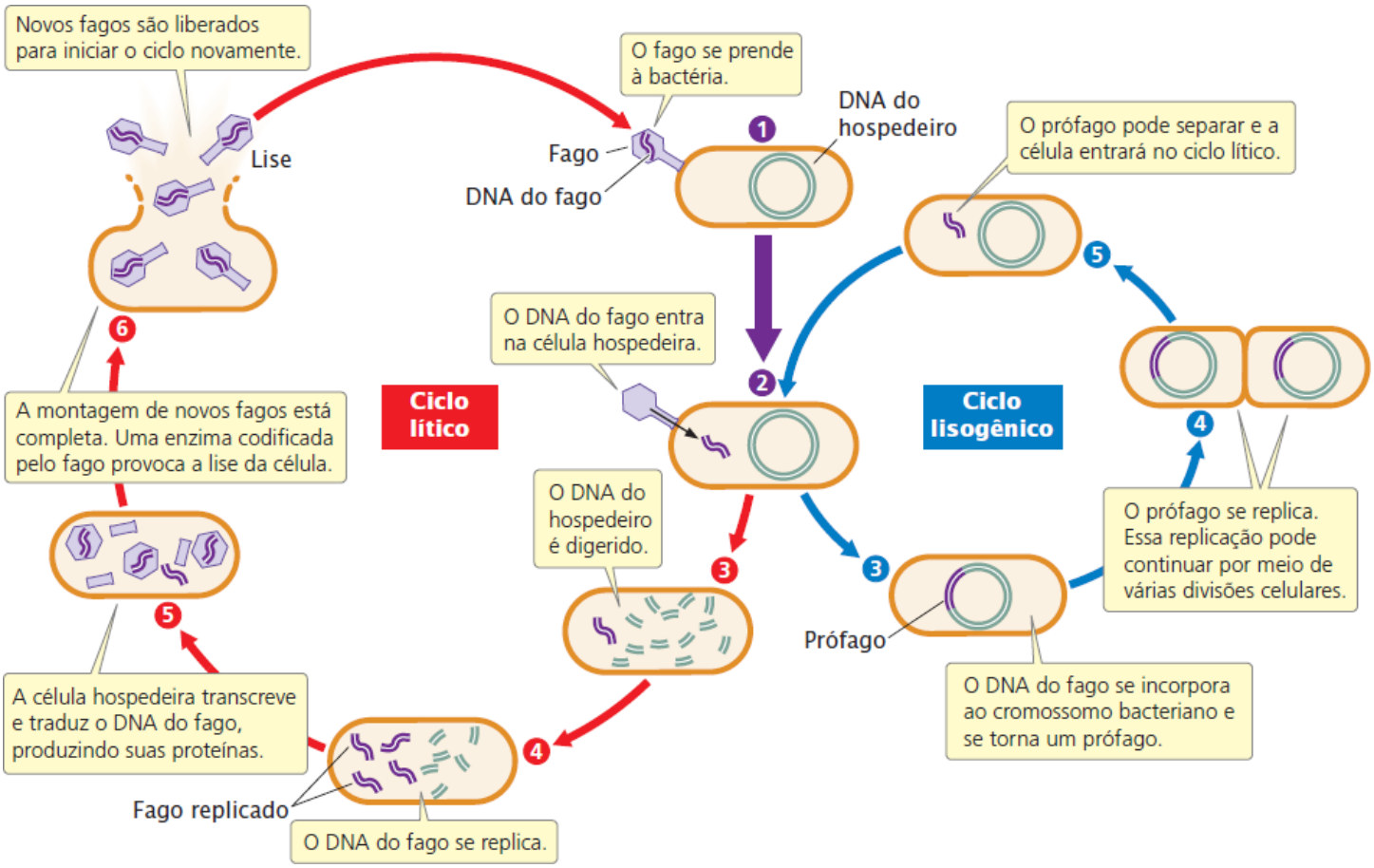


Figura 9.19 Os bacteriófagos têm dois ciclos de vida alternativos: os ciclos lítico e lisogênico.

Figura 9.20 As placas são manchas claras de células rompidas em um campo de bactérias. (Cortesia de Maria P. MacWilliams e Tong Lee.)

Conceitos

Os genomas dos vírus podem ser DNA ou RNA, circulares ou lineares e de fita dupla ou única. Os bacteriófagos são usados em vários tipos de pesquisa genética.

Checagem dos conceitos 6

Em qual ciclo de vida do bacteriófago o DNA do fago é incorporado ao cromossomo bacteriano?

a. Lítico.
b. Lisogênico.
c. Lítico e lisogênico.
d. Não é no lítico nem no lisogênico.

Transdução | Como usar os fagos para mapear os genes bacterianos

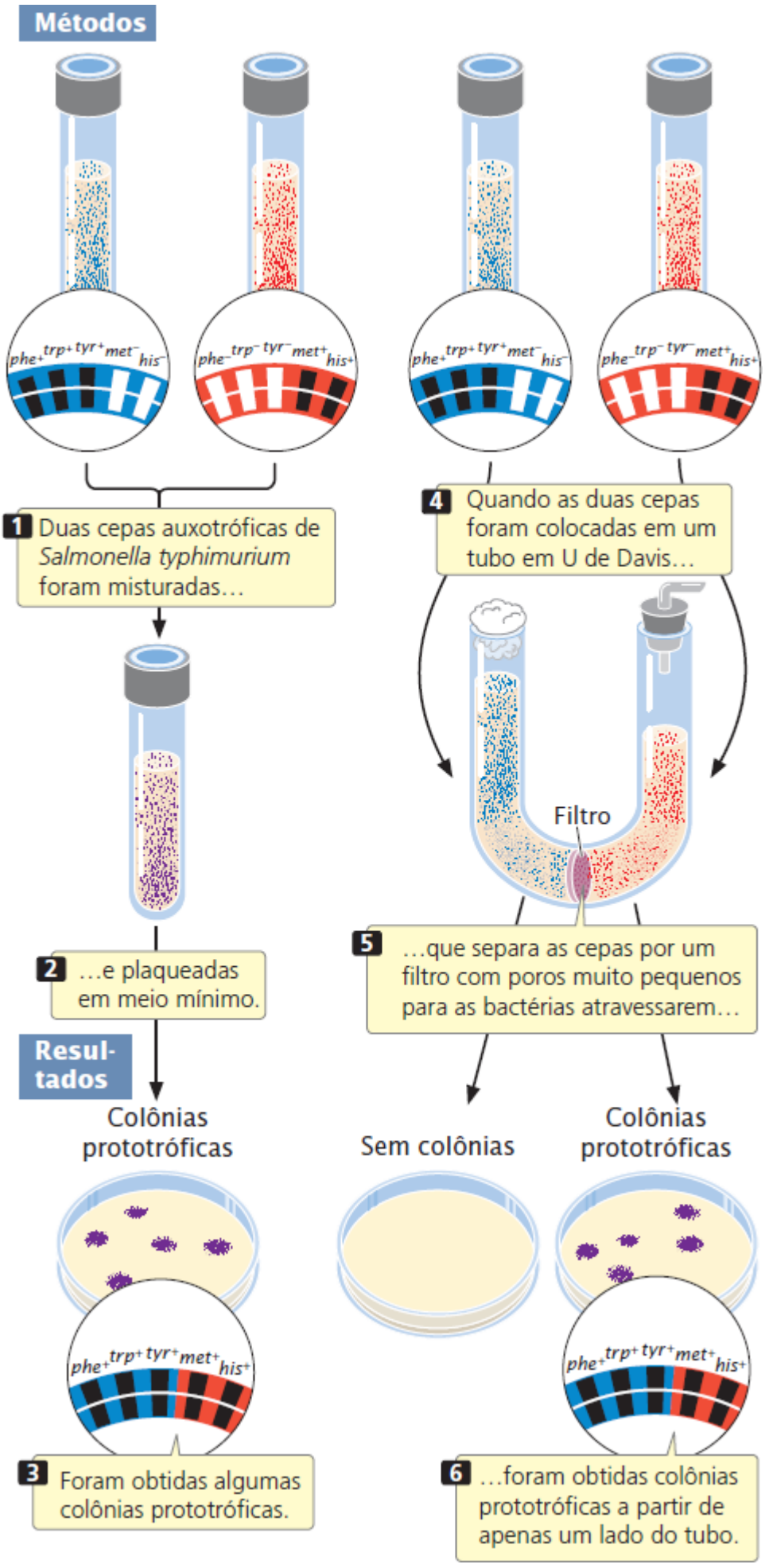
Na discussão sobre genética das bactérias, foram identificados três mecanismos de transferência genética: conjugação, transformação e transdução (ver Figura 9.6). Vamos examinar a transdução, na qual os genes são transferidos entre bactérias pelos vírus. Na **transdução generalizada**, qualquer gene pode ser transferido. Na **transdução especializada**, apenas alguns genes são transferidos.

Transdução generalizada. Joshua Lederberg e Norton Zinder descobriram a transdução generalizada em 1952, enquanto tentavam produzir recombinação na bactéria *Salmonella typhimurium* pela conjugação. Eles misturaram uma cepa de *S. typhimurium* que era *phe*$^+$ *trp*$^+$ *tyr*$^+$ *met*$^-$ *his*$^-$ com uma cepa que era *phe*$^-$ *trp*$^-$ *tyr*$^-$ *met*$^+$ *his*$^+$ (**Figura 9.21**) e as plaquearam em um meio mínimo. Surgiram alguns recombinantes prototróficos (*phe*$^+$ *trp*$^+$ *tyr*$^+$ *met*$^+$ *his*$^+$), sugerindo que ocorreu conjugação. Entretanto, quando testaram as duas cepas em um tubo em U semelhante ao usado por Davis, foram obtidos alguns prototróficos *phe*$^+$ *trp*$^+$ *tyr*$^+$ *met*$^+$ *his*$^+$ em um lado do tubo (compare a Figura 9.21 com a Figura 9.8). Esse aparato separou as duas cepas por um filtro com poros muito pequenos para a passagem das bactérias, então como os genes foram transferidos entre as bactérias

na ausência de conjugação? Os resultados de outros estudos revelaram que o agente da transferência era um bacteriófago.

Experimento

Pergunta: A troca genética entre bactérias precisa sempre do contato intercelular?



Conclusão: A troca genética não ocorreu por conjugação. Posteriormente, demonstrou-se que um fago foi o agente responsável pela transferência.

Figura 9.21 **O experimento de Lederberg e Zinder.**

No ciclo lítico da reprodução do fago, o cromossomo bacteriano é rompido em fragmentos aleatórios (**Figura 9.22**). Para alguns tipos de bacteriófagos, um pedaço do cromossomo bacteriano, em vez do DNA do fago, é empacotado no revestimento do fago; essas partículas de fago são chamadas **fagos transdutores**. O fago transdutor infecta uma nova célula, liberando o DNA bacteriano, e os genes introduzidos podem ser integrados ao cromossomo bacteriano por meio de um *crossing over* duplo. Alguns fagos transdutores inserem DNA viral, junto com o gene bacteriano, no cromossomo bacteriano. Em ambos os casos, os genes bacterianos podem ser deslocados de uma cepa para outra, produzindo bactérias recombinantes chamadas **transdutantes**.

Nem todos os fagos são capazes de fazer a transdução, um evento raro que requer (1) que o fago degrade o cromossomo bacteriano, (2) que o processo de empacotamento do DNA na proteína do fago não seja específico para o DNA do fago e (3) que os genes bacterianos transferidos pelo vírus se recombinem com o cromossomo na célula receptora. A taxa geral de transdução varia de apenas 1 em 100.000 a 1 em 1.000.000.

Em função do tamanho limitado de uma partícula de fago, somente cerca de 1% do cromossomo bacteriano pode ser transduzido. Apenas genes próximos no cromossomo bacteriano serão transferidos juntos ou **cotransduzidos**. Como a chance de uma célula ser transduzida por dois fagos separados é muito pequena, quaisquer genes cotransduzidos estão próximos no cromossomo bacteriano. Assim, as taxas de cotransdução, como as taxas de cotransformação, fornecem uma indicação das distâncias físicas entre os genes de um cromossomo bacteriano.

Para mapear genes usando a transdução são usadas duas cepas bacterianas com alelos diferentes em vários *loci*. A cepa doadora é infectada com fagos (**Figura 9.23**), que se reproduzem dentro da célula. Quando os fagos rompem as células doadoras, uma suspensão de descendentes do fago é misturada com uma cepa de bactérias receptoras, que então é plaqueada em diferentes tipos de meios para determinar os fenótipos dos fagos transdutores descendentes. Resolva o Problema 32

Conceitos

Na transdução, os genes bacterianos são compactados em um envoltório viral, transferidos para outra bactéria pelo vírus e incorporados no cromossomo bacteriano por meio do *crossing over*. Os genes bacterianos podem ser mapeados com o uso da transdução generalizada.

Checagem dos conceitos 7

Nos experimentos de mapeamento de genes usando a transdução generalizada, os genes bacterianos que são cotransduzidos estão:

a. Distantes no cromossomo bacteriano.
b. Em diferentes cromossomos bacterianos.
c. Próximos no cromossomo bacteriano.
d. Em um plasmídio.

Transdução especializada. Como a transdução generalizada, a transdução especializada requer a transferência de genes de uma bactéria para outra por meio dos fagos, mas aqui apenas os genes próximos de locais específicos no cromossomo bacteriano são transferidos. Esse processo requer bacteriófagos lisogênicos. O prófago pode ser retirado inadequadamente do cromossomo bacteriano, carreando com ele uma pequena parte do DNA bacteriano adjacente ao local da integração do prófago. Um fago carreando esse DNA o injeta em outra célula bacteriana no próximo ciclo de infecção. Esse processo lembra a situação das células F′, na qual o plasmídio F carreia os genes de uma bactéria para outra (ver Figura 9.12). A transdução especializada ocorre em fagos que utilizam locais específicos de integração no cromossomo bacteriano; muitos deles se integram de forma aleatória e exibem apenas a transdução generalizada.

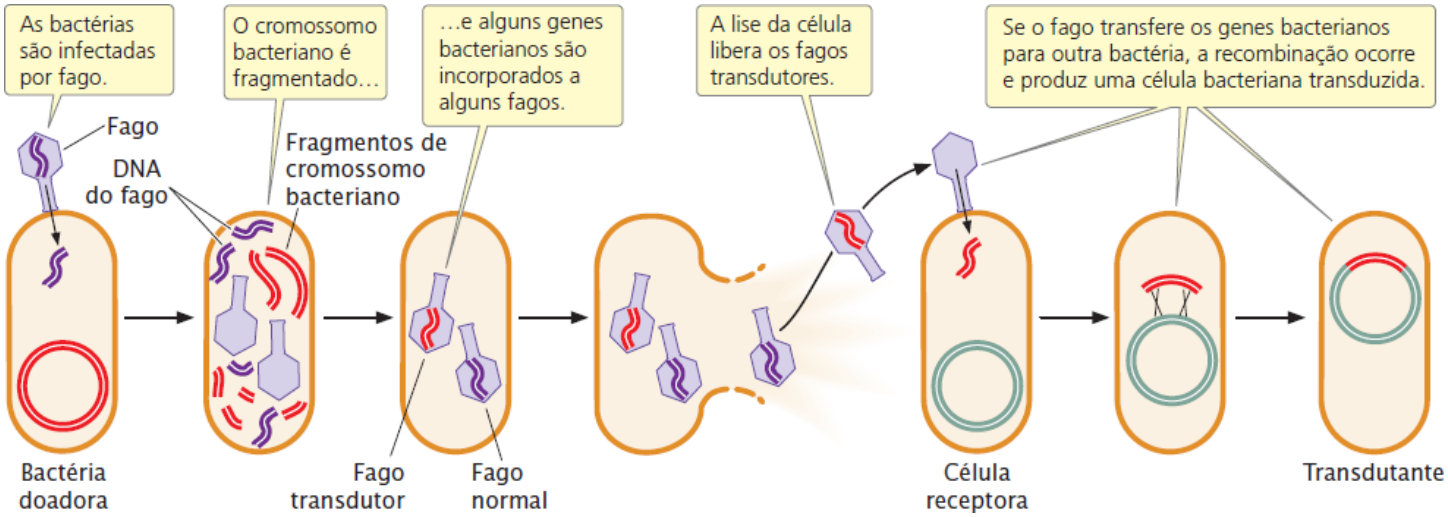


Figura 9.22 **Os genes podem ser transferidos de uma bactéria para outra por transdução generalizada.**

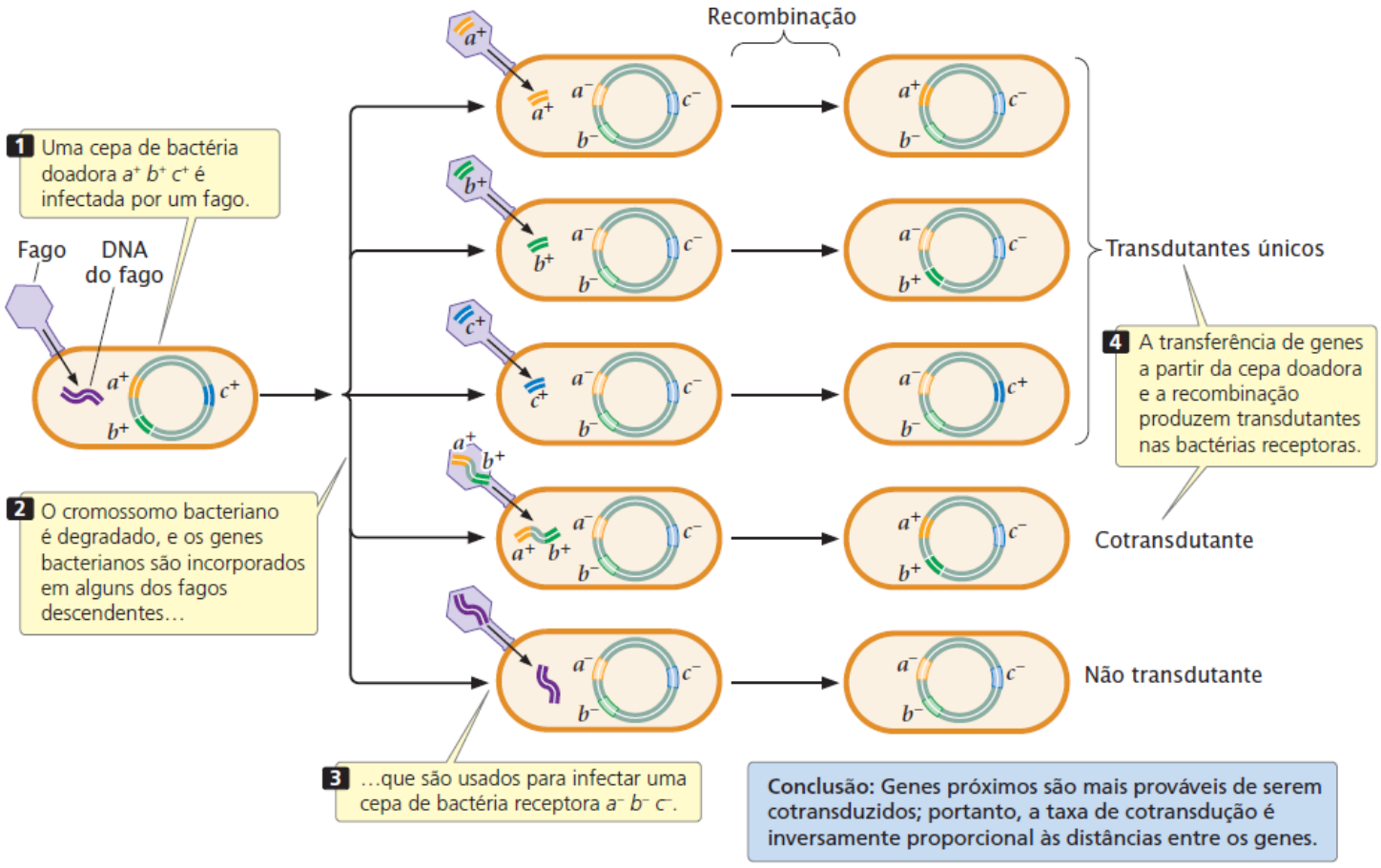


Figura 9.23 **A transdução generalizada pode ser usada para mapear os genes bacterianos.**

Conceitos

A transdução especializada transfere apenas os genes bacterianos localizados próximos ao local da inserção do prófago.

Conceitos conectantes

Três métodos para mapear os genes bacterianos

Foram destacados três métodos de mapeamento dos genes bacterianos: (1) conjugação interrompida; (2) transformação e (3) transdução. Esses métodos têm semelhanças e diferenças importantes.

O mapeamento por conjugação interrompida baseia-se no tempo necessário para os genes serem transferidos de uma bactéria para outra por contato intercelular. A chave para essa técnica é que o cromossomo bacteriano é transferido e a ordem dos genes e o tempo necessário para sua transferência

fornecem informações sobre as posições dos genes no cromossomo. Ao contrário dos outros métodos de mapeamento, a distância entre os genes é medida não nas frequências de recombinação, mas na unidade de tempo necessário para que os genes sejam transferidos. Aqui, a unidade básica do mapeamento por conjugação é um minuto.

No mapeamento dos genes por transformação, o DNA da cepa doadora é isolado, rompido e misturado com a cepa receptora. Alguns fragmentos passam para as células receptoras, onde o DNA transformado pode se recombinar com o cromossomo bacteriano. A unidade de transferência aqui é um fragmento aleatório do cromossomo. Os *loci* que estão próximos no cromossomo doado tendem a estar no mesmo fragmento de DNA; então, as taxas de cotransformação fornecem informações sobre as posições relativas dos genes no cromossomo.

O mapeamento por transdução também depende da transferência de genes entre as bactérias que têm um ou dois traços diferentes, mas aqui o veículo da transferência do gene é um bacteriófago. O mapeamento por transdução é semelhante àquele por transformação em vários aspectos. Pequenos fragmentos de DNA são transportados pelo fago a partir da bactéria doadora para a bactéria receptora, e as taxas de cotransdução, assim como as de cotransformação, fornecem informações sobre as distâncias relativas entre os genes.

Todos os métodos usam uma estratégia comum para mapear os genes bacterianos. O movimento dos genes da doadora para a receptora é detectado ao usar cepas com dois ou mais traços diferentes, e a transferência de um gene relativo à transferência dos outros é examinada. Além disso, todos os três métodos dependem da recombinação entre o DNA transferido e o cromossomo bacteriano. No mapeamento por conjugação interrompida, a ordem relativa e o momento da transferência do gene fornecem as informações necessárias para mapear os genes; na transformação e na transdução, a taxa de cotransferência fornece essas informações.

Ou seja, as mesmas estratégias básicas são usadas para o mapeamento por conjugação interrompida, transformação e transdução. Os métodos diferem principalmente nos seus mecanismos de transferência: no mapeamento por conjugação, o DNA é transferido por contato entre as bactérias; na transformação, o DNA é transferido como pequenos fragmentos descobertos; e na transdução, o DNA é transferido por bacteriófagos.

Mapeamento dos genes nos fagos

O mapeamento dos genes nos próprios bacteriófagos depende da recombinação homóloga entre os cromossomos do fago e, portanto, requer a aplicação dos mesmos princípios aos aplicados para mapear os genes nos organismos eucarióticos (ver *Capítulo 7*). São feitos cruzamentos entre vírus com dois ou mais genes diferentes, e os fagos descendentes recombinantes são identificados e contados. A proporção de descendentes recombinantes é então usada para estimar as distâncias entre os genes e sua ordem linear no cromossomo.

Em 1949, Alfred Hershey e Raquel Rotman examinaram as taxas de recombinação no bacteriófago T2, que tem uma única fita de DNA. Eles estudaram a recombinação entre genes em duas cepas com diferença na aparência da placa e faixa de hospedeiro (as cepas bacterianas que os fagos poderiam infectar). Uma cepa era capaz de infectar e romper as células *E. coli* tipo B, mas nem todas as células B/2 (tornando essa cepa de fato do tipo selvagem com a faixa do hospedeiro normal, ou h^+) e produziu uma placa anormal que era grande com bordas distintas (r^-). A outra cepa era capaz de infectar e romper *ambas* as células B *e* B/2 (faixa de hospedeiro mutante, h^-) e produzia placas do tipo selvagem que eram menores com bordas pouco definidas (r^+).

Hershey e Rotman cruzaram as cepas $h^+ r^-$ e $h^- r^+$ de T2 ao infectar as células *E. coli* do tipo B com uma mistura das duas cepas. Eles usaram uma concentração elevada de fagos de modo que a maioria das células seria infectada simultaneamente por ambas as cepas (**Figura 9.24**). Dentro das células bacterianas ocorreu recombinação homóloga entre os cromossomos de diferentes cepas de bacteriófagos, produzindo os cromossomos $h^+ r^+$ e $h^- r^-$, que eram então empacotados em novas partículas de fago. Quando as células rompiam, os fagos recombinantes eram liberados, junto com os fagos $h^+ r^-$ e $h^- r^+$ não recombinantes.

Hershey e Rotman diluíram e plaquearam os fagos descendentes em uma camada de bactérias composta por uma *mistura* das células B e B/2. Os fagos que carreavam o alelo h^+ (que conferia a capacidade de infectar apenas as células B) produziam uma placa turva porque as células B/2 não eram rompidas. Os fagos com o alelo h^- produziram uma placa clara porque todas as células dentro dessa placa eram rompidas. Os fagos r^+ produziram placas pequenas, enquanto os r^- produziram placas grandes. Os genótipos desses fagos descendentes poderiam ser determinados pela aparência da placa (ver Figura 9.24 e **Quadro 9.4**).

Nesse tipo de cruzamento de fago, a frequência de recombinação (*FR*) entre os dois genes pode ser calculada usando a fórmula a seguir:

$$FR = \frac{\text{placas recombinantes}}{\text{total de placas}}$$

No cruzamento de Hershey e Rotman, as placas recombinantes eram $h^+ r^+$ e $h^- r^-$; então a frequência de recombinação era

$$FR = \frac{(h^+ r^+) + (h^- r^-)}{\text{total de placas}}$$

As frequências de recombinação podem ser usadas para determinar as distâncias entre os genes e sua ordem no cromossomo do fago, como as frequências de recombinação são usadas para mapear os genes nos eucariotos. Resolva o Problema 35

Conceitos

Para mapear os genes dos fagos, as células bacterianas são infectadas com vírus com dois ou mais genes diferentes. As placas recombinantes são contadas e as taxas de recombinação são usadas para determinar a ordem linear dos genes no cromossomo e as distâncias entre eles.

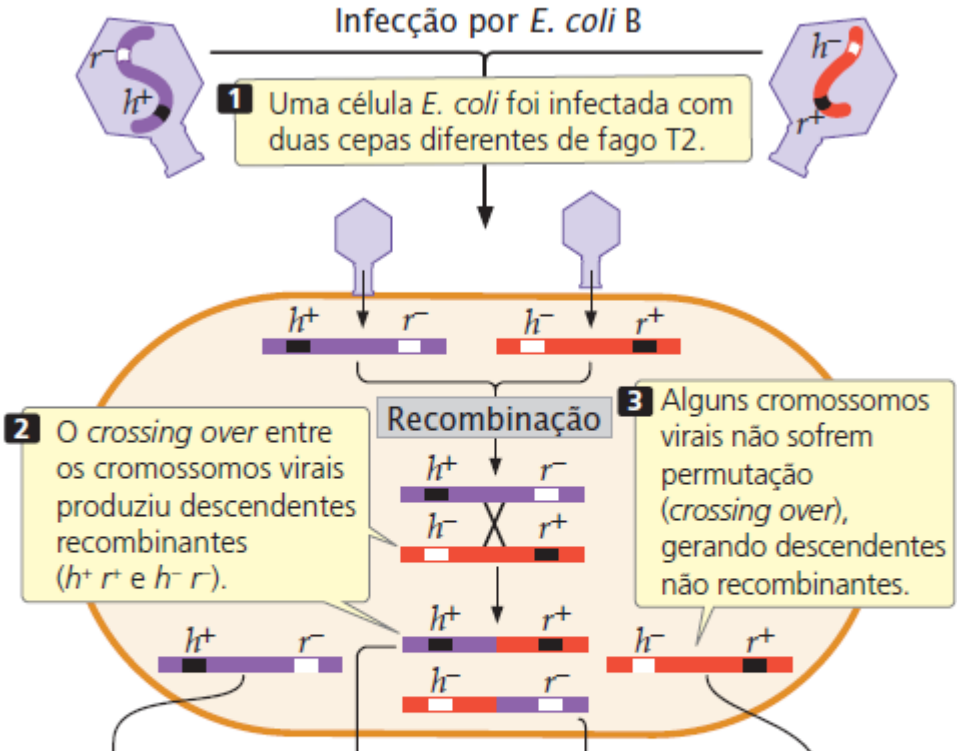
Quadro 9.4 Fagos descendentes produzidos a partir de $h^- r^+ \times h^+ r^-$.

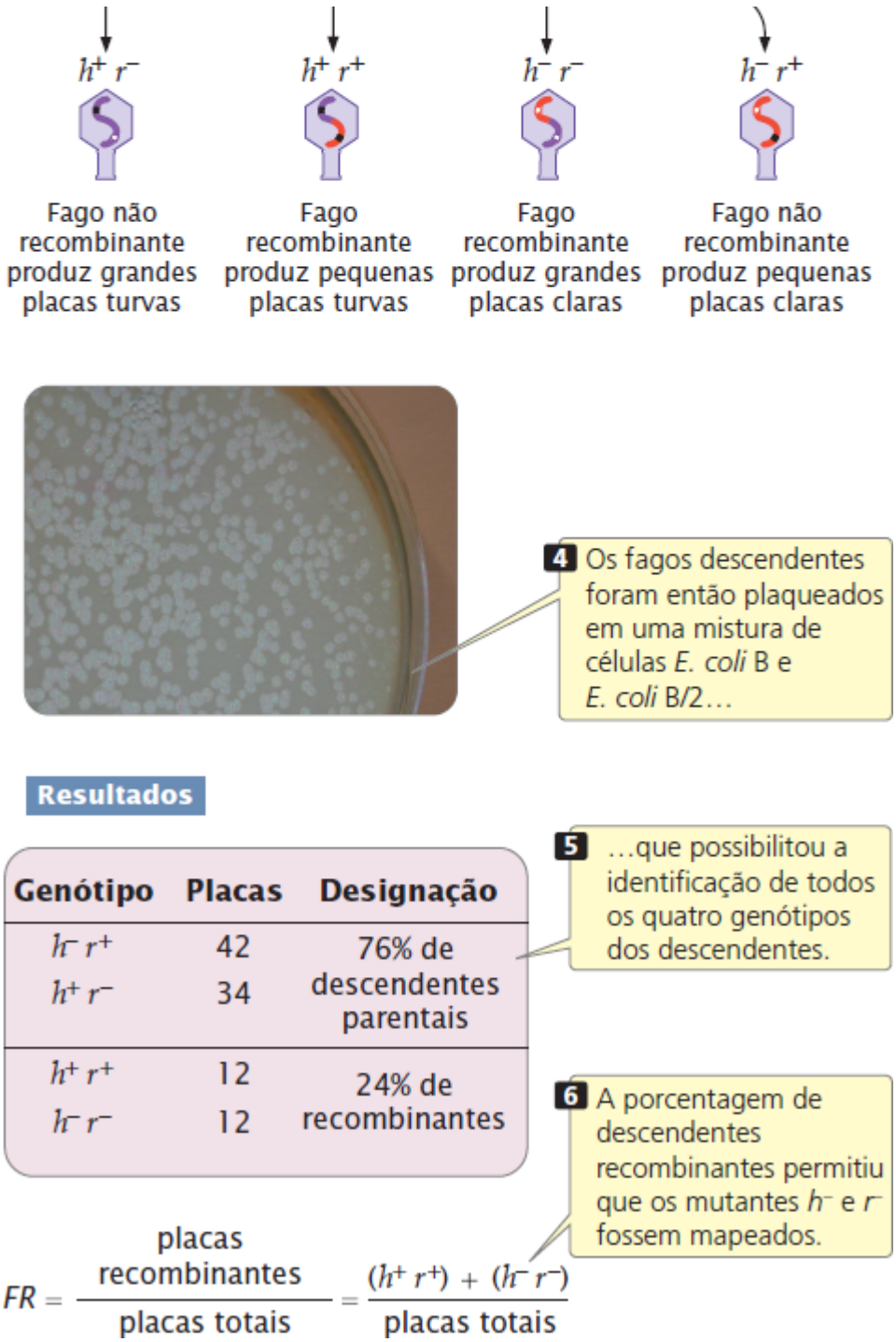
Fenótipo	Genótipo
Clara e pequena	$h^- r^+$
Turva e grande	$h^+ r^-$
Turva e pequena	$h^+ r^+$
Clara e grande	$h^- r^-$

Experimento

Pergunta: Como podemos determinar a posição de um gene em um cromossomo do fago?

Método





Conclusão: A frequência de recombinação indica que a distância entre os genes *h* e *r* é 24%.

Figura 9.24 Hershey e Rotman desenvolveram uma técnica para mapear os genes virais. (Cortesia de Steven R. Spilatro.)

Análise intraestrutural dos genes dos bacteriófagos

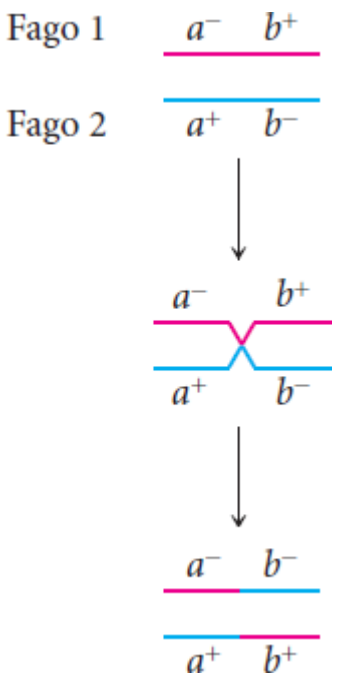
Nos anos de 1950 e 1960, Seymour Benzer conduziu diversos experimentos para examinar a estrutura de um gene. Como não havia técnicas moleculares naquele período para examinar diretamente as sequências de nucleotídios, Benzer teve que deduzir a estrutura do gene para análises das mutações e seus efeitos. Os resultados de seus estudos mostraram que diferentes locais de mutação *dentro* de um único gene poderiam ser mapeados (chamado de **mapeamento intragênico**) ao usar técnicas semelhantes às descritas para mapear os genes bacterianos por transdução. Diferentes locais de um único gene estão muito próximos, então a recombinação entre eles ocorre em uma frequência muito baixa. Como são necessários vários descendentes para detectar essas recombinações, Benzer usou o bacteriófago T4, que se reproduz rapidamente e produz muitos descendentes.

Técnicas de mapeamento de Benzer. Os fagos T4 do tipo selvagem produzem placas pequenas com bordas rugosas quando crescem em uma camada de *E. coli*. Alguns mutantes, chamados de *r* por causa da lise rápida, produzem placas maiores com bordas bem-definidas. Benzer isolou fagos com várias mutações *r* diferentes, concentrando-se em um subgrupo particular

chamado de mutantes *rII*.

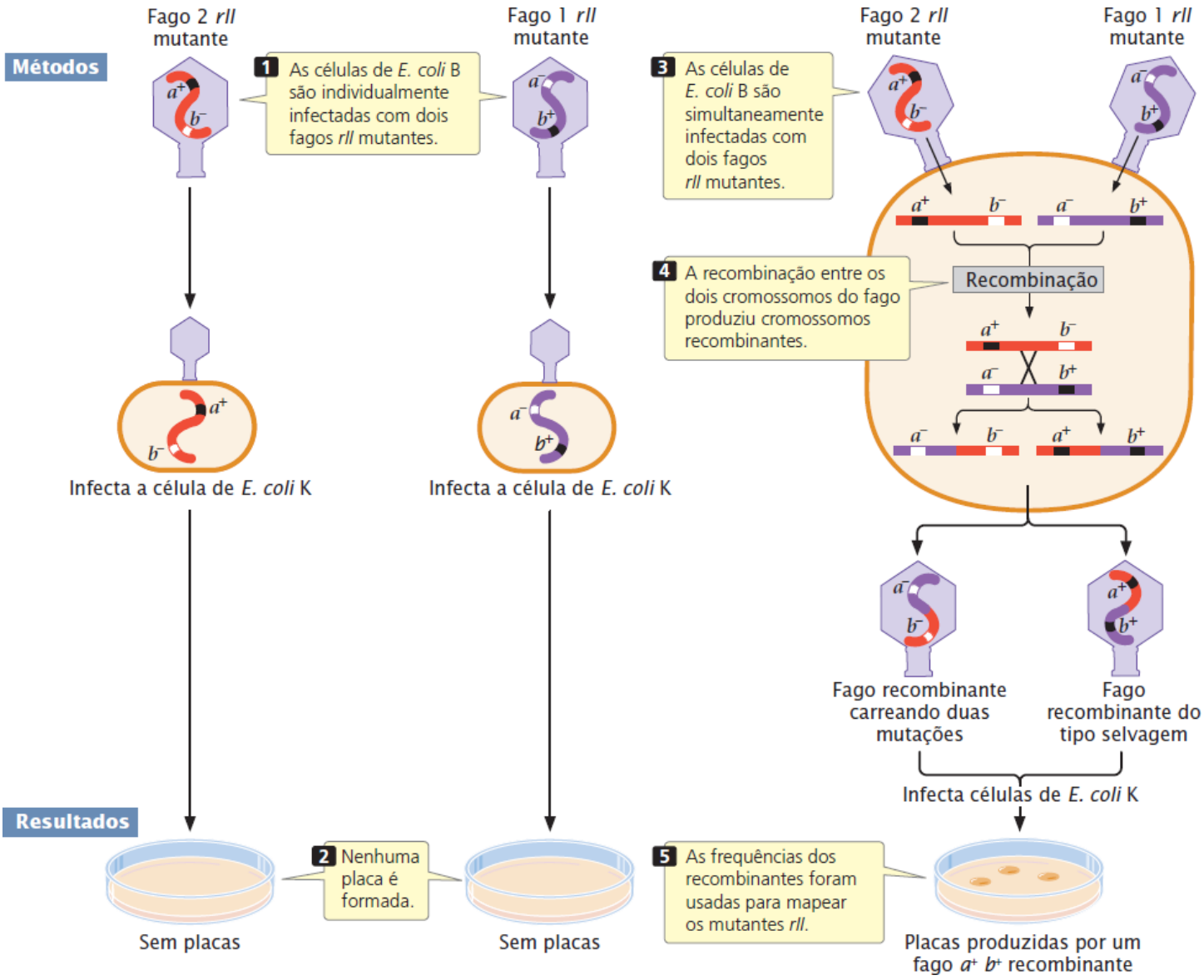
Os fagos T4 do tipo selvagem produzem placas típicas nas cepas B e K de *E. coli*. Ao contrário, os mutantes *rII* produzem placas na cepa B e não formam placas na cepa K. Benzer identificou os mutantes *r* por meio das suas placas diferentes quando cresciam na *E. coli* B. Ele coletou o lisado dessas placas e usou para infectar *E. coli* K. Os fagos que não produziam placas na *E. coli* K eram definidos como o tipo *rII*.

Benzer coletou milhares de mutações *rII*. Ele infectou simultaneamente células bacterianas com dois mutantes diferentes e procurou por descendentes recombinantes (**Figura 9.25**). Considere duas mutações *rII*, a^- e b^- (seus alelos do tipo selvagem são a^+ e b^+). Benzer infectou células de *E. coli* B com duas cepas diferentes de fagos, uma $a^-\ b^+$ e outra $a^+\ b^-$ (ver Figura 9.25, etapa 3). Nenhuma dessas mutações foi capaz de crescer nas *E. coli* K. Enquanto se reproduziam nas cepas B, alguns fagos das duas cepas se recombinaram (ver Figura 9.25, etapa 4). Um único *crossing over* produz dois cromossomos recombinantes, um com genótipo $a^+\ b^+$ e outro com genótipo $a^-\ b^-$:



Experimento

Pergunta: Como os fagos *rII* mutantes podem ser mapeados e o que eles podem revelar sobre a estrutura do gene?



Conclusão: O mapeamento de mais de 2.400 mutantes *rII* forneceu informações sobre a estrutura interna de um gene em nível de par de base – a primeira visão da estrutura molecular de um gene.

Figura 9.25 Benzer desenvolveu um procedimento para mapear os mutantes *rII*. Dois mutantes *rII* diferentes ($a^-\ b^+$ e $a^+\ b^-$) são isolados de células de *E. coli* B. Apenas o recombinante $a^+\ b^+$ consegue crescer na célula da *E. coli* K, possibilitando a identificação desses recombinantes.

Os cromossomos recombinantes resultantes, junto com os cromossomos não recombinantes (parentais) foram incorporados nos fagos descendentes, que então foram usados para infectar *E. coli* K. As placas resultantes foram examinadas para determinar o genótipo do fago infectante e mapear os mutantes *rII* (ver Figura 9.25, etapa 5).

Nenhum dos mutantes *rII* cresceu em *E. coli* K (ver Figura 9.25, etapa 2), mas os fagos do tipo selvagem cresceram; então, os fagos descendentes que produziram placas nas cepas K devem ter o genótipo recombinante $a^+\ b^+$. Cada recombinação produz um número igual de mutantes duplos ($a^-\ b^-$) e cromossomos do tipo selvagem ($a^+\ b^+$), de modo que o número de descendentes recombinantes deve ser o dobro do número de placas do tipo selvagem que aparecem na cepa K. A frequência de recombinação entre os dois mutantes *rII* seria:

$$\text{FR} = \frac{2 \times \text{número de placas na } \textit{E. coli} \text{ K}}{\text{número total de placas na } \textit{E. coli} \text{ B}}$$

Como os fagos produzem muitos descendentes, Benzer pode detectar um único recombinante entre bilhões de fagos descendentes. As frequências de recombinação são proporcionais às distâncias físicas ao longo do cromossomo (ver *Capítulo 7*), revelando as posições em diferentes mutações dentro da região *rII* do cromossomo do fago. Assim, Benzer conseguiu mapear mais de 2.400 mutações *rII*, muitas correspondendo a pares de bases únicas no DNA viral. Seu trabalho forneceu a

primeira visão molecular de um gene.

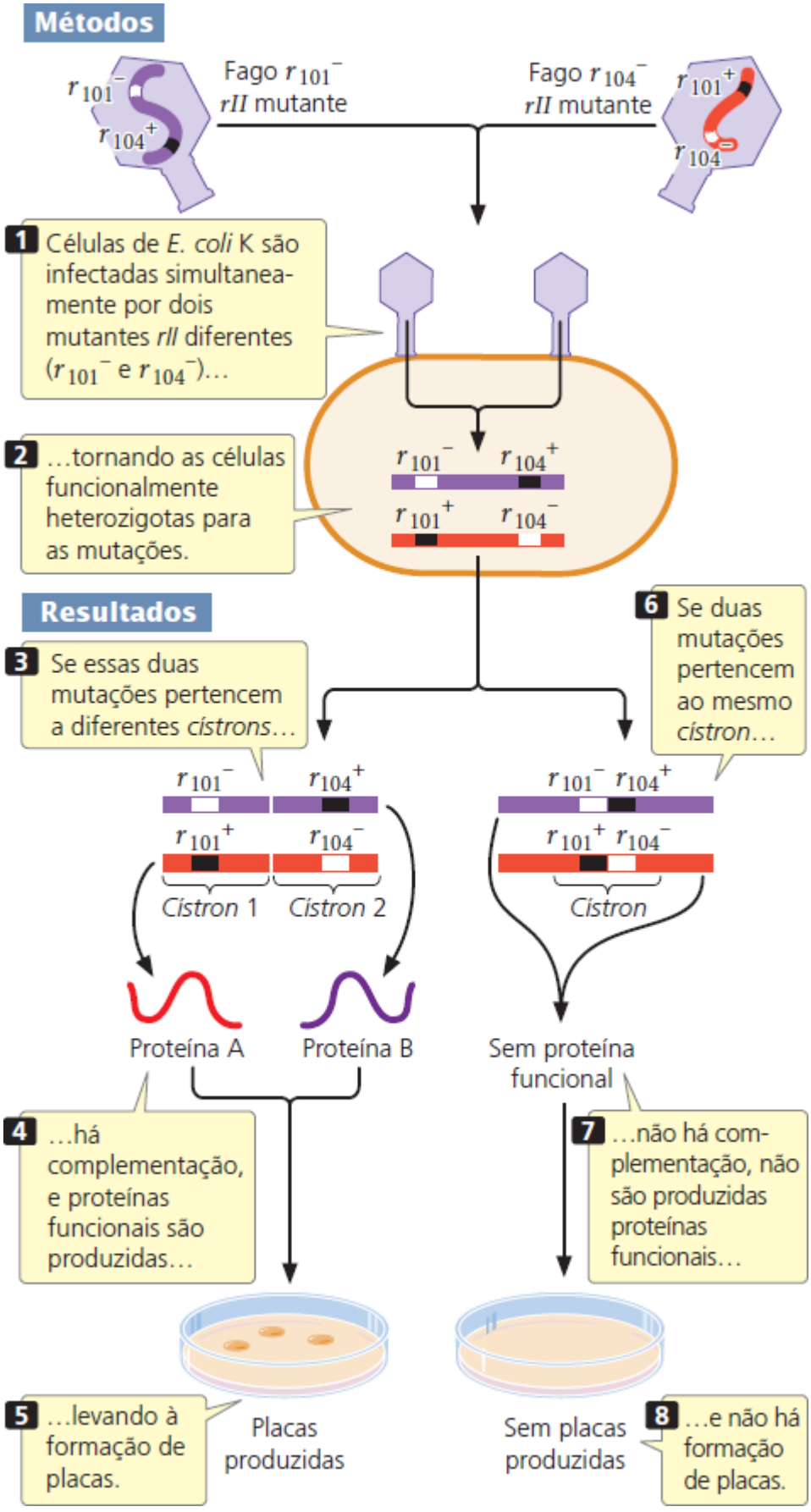
Experimentos de complementação. Os experimentos de mapeamento de Benzer demonstraram que algumas mutações *rII* tinham uma ligação muito próxima. Esse achado levantou a questão se elas estavam no mesmo *locus* ou em *loci* diferentes. Para determinar se diferentes mutações *rII* pertenciam a *loci* funcionais diferentes, Benzer usou o teste de complementação (*cis-trans*) (ver *Capítulo 5*).

Para o teste de complementação no bacteriófago, Benzer infectou cepas K com muitos fagos de duas cepas mutantes do fago (**Figura 9.26**, etapa 1) de modo que as células seriam duplamente infectadas por ambas as cepas. Considere duas mutações *rII*: r_{101}^{-} e r_{104}^{-}. As cepas infectadas pelos dois mutantes

$$\frac{r_{101}^{-}\ r_{104}^{+}}{r_{101}^{+}\ r_{104}^{-}}$$

Experimento

Pergunta: Como determinamos se duas mutações *rII* diferentes estão no mesmo locus?



Conclusão: O teste de complementação indica se as duas mutações estão no mesmo locus ou em diferentes *loci.*

Figura 9.26 Os testes de complementação são usados para determinar se diferentes mutações estão no mesmo gene funcional.

eram efetivamente heterozigotas para os genes do fago, com as mutações na configuração *trans* (ver Figura 9.26, etapa 2). No

teste de complementação, os fenótipos dos fagos descendentes eram examinados na cepa K em vez da cepa B, como ilustrado na Figura 9.25.

Se as mutações r_{101}^- e r_{104}^- estiverem em diferentes *loci* funcionais que codificam diferentes proteínas, então nas células bacterianas infectadas por ambos mutantes as sequências do tipo selvagem no cromossomo oposto de cada mutação sobrepõem os efeitos das mutações recessivas; os fagos produzirão placas normais nas cepas K (Figura 9.26, etapas 3, 4 e 5). Se, por outro lado, as mutações estiverem no mesmo *locus*, nenhuma proteína funcional é produzida por um ou outro cromossomo e não surgem placas nas células K (Figuras 9.26, etapas 6, 7 e 8). Assim, a ausência de placas indica que as duas mutações estão no mesmo *locus*. Benzer criou o termo *cístron* para indicar um gene funcional definido pelo teste de complementação.

No teste de complementação, o heterozigoto *cis* é usado como controle. Benzer infectou simultaneamente bactérias com fagos do tipo selvagem ($\underline{r_{101}^+ r_{104}^+}$) e com fagos que carreavam ambas as mutações ($\underline{r_{101}^- r_{104}^-}$). Esse teste produziu células que eram heterozigotas, e na configuração *cis* para os genes do fago:

$$\frac{r_{101}^+ \; r_{104}^+}{r_{101}^- \; r_{104}^-}$$

Independente se as mutações r_{101}^- e r_{104}^- estão na mesma unidade funcional, essas células têm uma cópia do cromossomo do fago do tipo selvagem ($\underline{r_{101}^+ r_{104}^+}$) e produzirão placas normais na célula de *E. coli* K.

Benzer fez o teste de complementação em muitos pares de mutantes *rII*. Ele descobriu que a região *rII* tinha dois *loci*, indicados como *rIIA* e *rIIB*. As mutações nos grupos *rIIA* e *rIIB* se complementavam, mas as mutações no grupo *rIIA* não se complementavam com outras na *rIIA*; nem as do grupo *rIIB* se complementavam com outras na *rIIB*.

Conceitos

Em uma série de experimentos com o bacteriófago T4, Seymour Benzer mostrou que a recombinação poderia ocorrer em um único gene e criou o primeiro mapa molecular de um gene. Ele usou o teste de complementação para distinguir entre genes funcionais (*loci*).

Checagem dos conceitos 8

Nos testes de complementação, Benzer infectou simultaneamente as células *E. coli* com dois fagos, cada um com uma mutação diferente. Que conclusão ele tirou quando os fagos descendentes produziram placas normais?

a. As mutações estavam no mesmo *locus*.
b. As mutações estavam em *loci* diferentes.
c. As mutações estavam próximas no cromossomo.
d. Os genes estavam na configuração *cis*.

A relação entre os genes e a estrutura do DNA era desconhecida no período da pesquisa de Benzer. Um gene era definido como uma unidade funcional de hereditariedade que codificava um fenótipo. Muitos geneticistas acreditavam que os genes eram indivisíveis e que não aconteceria recombinação entre eles. Benzer demonstrou que a recombinação intragênica acontecia (embora em uma taxa muito baixa) e deu aos geneticistas a primeira luz sobre a estrutura de um gene. Resolva o Problema 40

Vírus com RNA

Até o momento, consideramos os vírus que infectam principalmente as bactérias. Os vírus também infectam plantas e animais, e alguns são patógenos importantes para esses organismos. O que aprendemos sobre os bacteriófagos tem implicações importantes para os vírus que infectam os organismos mais complexos.

Os genomas virais podem estar codificados em DNA ou RNA. O RNA é o material genético de alguns vírus humanos, incluindo os que causam resfriados, *influenza*, poliomielite e AIDS. Quase todos os vírus que infectam plantas têm genomas de RNA. A importância médica e econômica dos vírus RNA tem estimulado seu estudo.

Os vírus RNA capazes de se integrar aos genomas de seus hospedeiros, como fagos temperados se inserem no cromossomo das bactérias, são chamados de **retrovírus** (**Figura 9.27 A**). Como o genoma retroviral é o RNA, enquanto o hospedeiro é DNA, um retrovírus precisa produzir a **transcriptase reversa**, uma enzima que sintetiza DNA complementar (cDNA) a partir de um RNA ou um molde de DNA. Um retrovírus usa a transcriptase reversa para copiar seu genoma de RNA em uma

molécula de DNA de fita única, e a enzima transcriptase reversa – ou, às vezes, a DNA polimerase do hospedeiro – copia esse DNA de fita única, criando uma molécula de DNA de fita dupla. A cópia de DNA do genoma viral se integra ao cromossomo do hospedeiro. Um genoma viral incorporado ao cromossomo do hospedeiro é chamado de **pró-vírus**. O pró-vírus é replicado pelas enzimas do hospedeiro quando o cromossomo do hospedeiro é replicado (**Figura 9.27 B**).

Quando as condições são adequadas, o pró-vírus sofre transcrição para produzir várias cópias do genoma RNA original do vírus. Esse RNA codifica as proteínas virais e serve como RNA genômico para novas partículas virais. À medida que os vírus saem da célula, eles coletam partes da membrana celular para usar como seus envelopes.

Todos os genomas retrovirais conhecidos têm três genes em comum: *gag*, *pol* e *env*, cada um codificando uma proteína precursora que é clivada em duas ou mais proteínas funcionais. O gene *gag* codifica proteínas que constituem o revestimento proteico do vírus. O gene *pol* codifica a transcriptase reversa e uma enzima chamada **integrase**, que insere o DNA viral no cromossomo do hospedeiro. O gene *env* codifica as glicoproteínas que aparecem na superfície do vírus.

Alguns retrovírus têm **oncogenes** (ver *Capítulo 23*) que podem estimular a divisão celular e levar a formação de tumores. O primeiro retrovírus a ser isolado, o vírus do sarcoma de Rous, foi identificado originalmente por sua capacidade de produzir tumores no tecido conjuntivo (sarcomas) nas galinhas.

Vírus da imunodeficiência humana e AIDS

Um exemplo de retrovírus é o vírus da imunodeficiência humana (HIV), que causa a síndrome da imunodeficiência adquirida (AIDS). A AIDS foi identificada pela primeira vez em 1982, quando vários homens homossexuais nos EUA começaram a apresentar sintomas de uma nova doença de deficiência do sistema imunológico. Naquele ano, Robert Gallo propôs que a AIDS era causada por um retrovírus. Entre 1983 e 1984, à medida que a epidemia da AIDS se disseminava, o retrovírus HIV foi isolado em pacientes com AIDS. Atualmente, sabemos que a AIDS é causada por dois vírus diferentes da imunodeficiência, HIV-1 e HIV-2, que juntos infectaram mais de 60 milhões de pessoas no mundo. Desses, 30 milhões morreram. A maioria dos casos de AIDS é causada pelo HIV-1, que tem distribuição global; o HIV-2 é encontrado principalmente na África Ocidental.

O HIV ilustra a importância da recombinação genética na evolução do vírus. Estudos das sequências de DNA do HIV e de outros retrovírus mostraram que o HIV-1 está intimamente relacionado com o vírus da imunodeficiência símia, encontrado em chimpanzés (SIV_{cpz}). Muitos chimpanzés selvagens na África são infectados pelo SIV_{cpz}, embora ele não provoque os sintomas como a AIDS nesses mamíferos. O SIV_{cpz} é, por si só, um híbrido resultante da recombinação entre um retrovírus encontrado no mangabei-de-penacho-vermelho (um macaco) e um retrovírus encontrado no macaco-de-nariz-branco (**Figura 9.28**). Aparentemente, um ou mais chimpanzés foram infectados por ambos os vírus; a recombinação entre eles produziu o SIV_{cpz}, que foi então transmitido para os seres humanos pelo contato com chimpanzés infectados. Nos seres humanos, o SIV_{cpz} sofreu uma importante evolução para se tornar o HIV-1, que então se espalhou pelo mundo, causando epidemia da AIDS. Várias transferências independentes de SIV_{cpz} para os seres humanos deram origem a diferentes cepas de HIV-1. O HIV-2 evoluiu a partir de um retrovírus diferente, SIV_{sm}, encontrado nos mangabeis cinzentos.

O HIV é transmitido pelo contato sexual entre humanos e por qualquer tipo de contato sangue a sangue, como o compartilhamento de agulhas contaminadas por dependentes químicos. Ele também pode ser transmitido da mãe para filho durante a gravidez e pelo aleitamento. Até que os testes de rastreamento pudessem identificar o sangue infectado pelo HIV, as transfusões e os fatores de coagulação usados pelos hemofílicos também eram fontes de infecção.

O HIV ataca principalmente uma classe de células sanguíneas chamadas de linfócitos T ou, simplesmente, células T auxiliares. O vírus entra em uma célula T auxiliar, sofre transcrição reversa e se integra ao cromossomo. O HIV se reproduz rapidamente, destruindo a célula T à medida que novas partículas de vírus saem da célula. Como as células T auxiliares são cruciais para a função imunológica e são destruídas pela infecção, os pacientes com AIDS têm uma resposta imunológica reduzida. A maioria dos pacientes com AIDS morre por causa de infecções secundárias que desenvolvem porque perderam a capacidade de lutar contra os patógenos.

O genoma do HIV tem 9.749 nucleotídios e carreia os genes *gag*, *pol*, *env* e outros seis que regulam o ciclo de vida do vírus. A transcriptase reversa do HIV é muito suscetível a erros, dando ao vírus uma elevada taxa de mutação e permitindo que ele evolua rapidamente, mesmo em um único hospedeiro. Essa rápida evolução dificulta o desenvolvimento de uma vacina eficiente contra o HIV. A variação genética dentro da população humana também afeta o vírus. Até o momento foram identificados mais de 10 *loci* nos seres humanos que afetam a infecção pelo HIV e o progresso da AIDS.

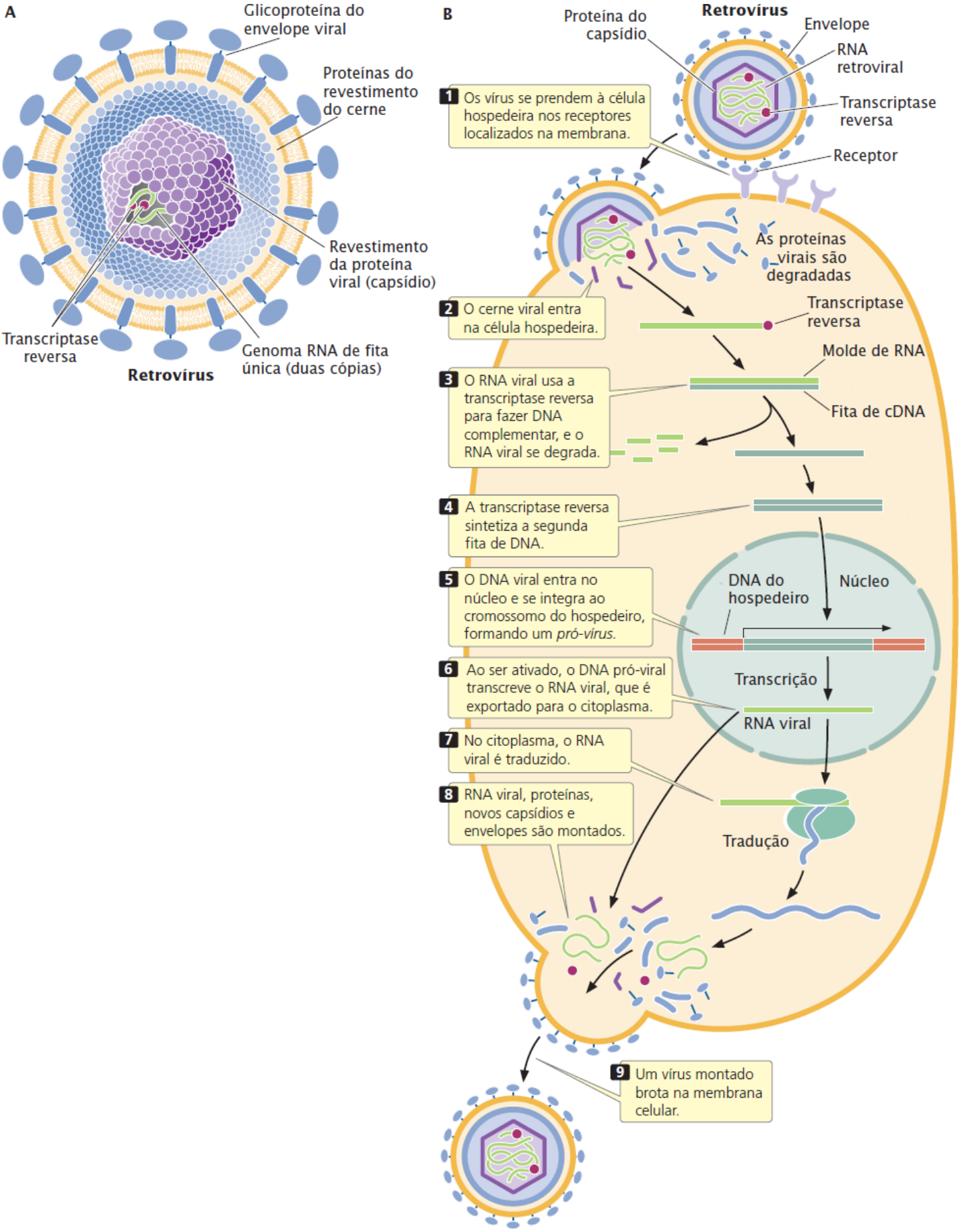


Figura 9.27 Um retrovírus usa a transcrição reversa para incorporar seu RNA ao DNA hospedeiro. A. Estrutura de um retrovírus típico. Duas cópias de um genoma de RNA de fita única e a enzima transcriptase reversa são mostradas envoltas por um capsídio de proteína. O capsídio é circundado por um envelope viral recoberto com glicoproteínas virais. **B.** Ciclo de vida de um retrovírus.

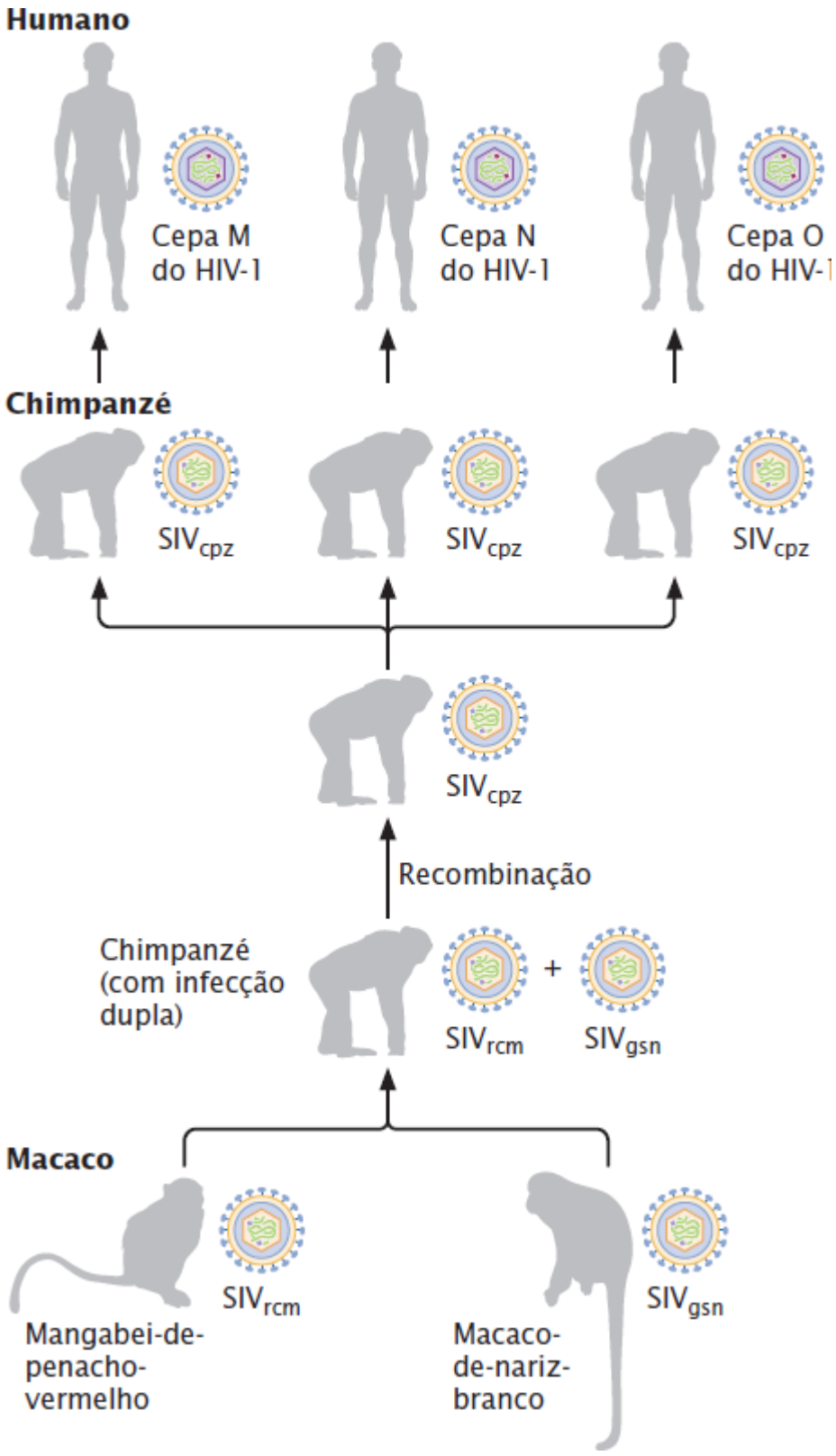


Figura 9.28 O HIV-1 evoluiu a partir de um vírus semelhante (SIV_{cpz}) encontrado nos chipanzés e foi transmitido para os seres humanos. O SIV_{cpz} surgiu a partir da recombinação ocorrida entre o retrovírus de mangabeis-de-penacho-vermelho e de macacos-de-nariz-branco.

Conceitos

Um retrovírus é um vírus RNA que se integra ao cromossomo do hospedeiro ao fazer uma cópia do DNA do seu genoma de RNA por meio do processo de transcrição reversa. O vírus da imunodeficiência humana, agente causador da AIDS, é um retrovírus. Ele evoluiu a partir de retrovírus próximos encontrados em outros primatas.

Checagem dos conceitos 9

Que enzima é usada por um retrovírus para fazer uma cópia de DNA de seu genoma?

Influenza | Gripe

A *influenza* é a prova de como as rápidas mudanças em um patógeno podem surgir a partir da recombinação do seu material genético. Também chamada de gripe, esta é uma doença respiratória causada pelos vírus influenza. Nos EUA, de 5 a 20% da população contraem *influenza* todos os anos e, embora a maioria dos casos seja leve, estima-se que 36.000 pessoas morram por causa dessa infecção anualmente. Em algum momento, em particular quando novas cepas do vírus influenza entram na população humana, ocorre uma epidemia mundial (chamada de pandemia). Em 1918, por exemplo, o vírus da gripe espanhola matou entre 20 milhões e 100 milhões de pessoas no mundo.

Os vírus influenza são vírus RNA que infectam aves e mamíferos. Os três tipos principais são influenza A, influenza B e influenza C. A maioria de casos de gripe comum é causada pelos vírus A e B. O vírus influenza A é dividido em subtipos com base nas duas proteínas, hemaglutinina (HA) e neuraminidase (NA), encontradas na superfície do vírus. As proteínas HA e NA afetam a capacidade do vírus em invadir as células hospedeiras e a resposta do sistema imunológico do organismo hospedeiro em responder à infecção. Existem 16 tipos de HA e 9 tipos de NA, que podem existir em um vírus em diferentes combinações. Por exemplo, cepas comuns de vírus influenza A que circulam nos seres humanos são H1N1 e H3N2 (**Quadro 9.5**), junto com várias cepas de vírus influenza B. A maioria dos subtipos diferentes de influenza A é encontrada nas aves.

Embora o influenza seja um vírus RNA, ele não é um retrovírus: seu genoma não é copiado para o DNA e incorporado ao cromossomo do hospedeiro como ocorre com um retrovírus. O genoma viral do vírus influenza tem sete a oito segmentos de RNA que são encapsulados em um envelope viral. Cada segmento de RNA codifica uma ou duas proteínas do vírus. O vírus invade a célula hospedeira ao se prender a receptores específicos na membrana celular. Após a partícula viral ter entrado na célula, o RNA viral é liberado, copiado e traduzido em proteínas virais. As moléculas de RNA e as proteínas virais são então montadas em novas partículas virais, que deixam a célula e infectam outras células.

Um dos perigos do vírus influenza é que ele evolui rapidamente, com novas cepas aparecendo frequentemente. A *influenza* evolui de duas formas. Primeiro, cada cepa muda continuamente por meio de mutações que surgem no RNA viral. A enzima que copia o RNA é especialmente suscetível a erros; então, novas mutações são introduzidas de forma contínua no genoma viral. Esse tipo de mudança contínua é chamado **variação antigênica**. Às vezes, as grandes mudanças no genoma viral ocorrem por meio da **mudança antigênica**, com o material genético de diferentes cepas sendo combinado em um processo chamado recombinação. A recombinação ocorre quando um hospedeiro é infectado simultaneamente com duas cepas diferentes. Os RNAs de ambas as cepas são replicados dentro da célula, e os segmentos de DNA das duas cepas são incorporados na mesma partícula viral criando uma nova cepa. Por exemplo, em 2002 ocorreu uma recombinação entre os subtipos H1N1 e H3N2, criando uma nova cepa H1N2 com a hemaglutinina do H1N1 e a neuraminidase do H3N2. As novas cepas produzidas pela mudança antigênica são responsáveis pela maioria das pandemias, porque ninguém tem imunidade a vírus tão diferentes produzidos.

Quadro 9.5 Cepas do vírus influenza responsáveis pelas principais pandemias de gripe.

Ano	Pandemia da gripe	Cepa
1918	Gripe espanhola	H1N1
1957	Gripe asiática	H2N2
1968	Gripe de Hong Kong	H3N2
2009	Gripe suína	H1N1

A maioria das cepas diferentes de influenza A infecta as aves. Os seres humanos não são infectados com muita facilidade pela *influenza* aviária. O surgimento de novas cepas nos seres humanos é causado pelos vírus que se recombinam nos porcos, que podem ser infectados pelos vírus dos tipos humano e aviário. Em 2009, uma nova cepa de vírus influenza H1N1 (chamada de gripe suína) surgiu no México e se espalhou rapidamente pelo mundo. Esse vírus surgiu a partir de uma série de eventos de recombinação que arranjou sequências de genes de vírus influenza humano, aviário e suíno e produziu um novo vírus H1N1 (**Figura 9.29**). Os hábitos dos fazendeiros que criam porcos e aves em locais próximos uns dos outros podem ter facilitado a recombinação entre as cepas aviária, suína e humana de influenza.

Conceitos

A *influenza* é causada por vírus RNA da influenza. Novas cepas de vírus influenza surgem por meio da mudança antigênica, na qual novos genomas virais são criados por recombinação de moléculas de RNA de diferentes cepas.

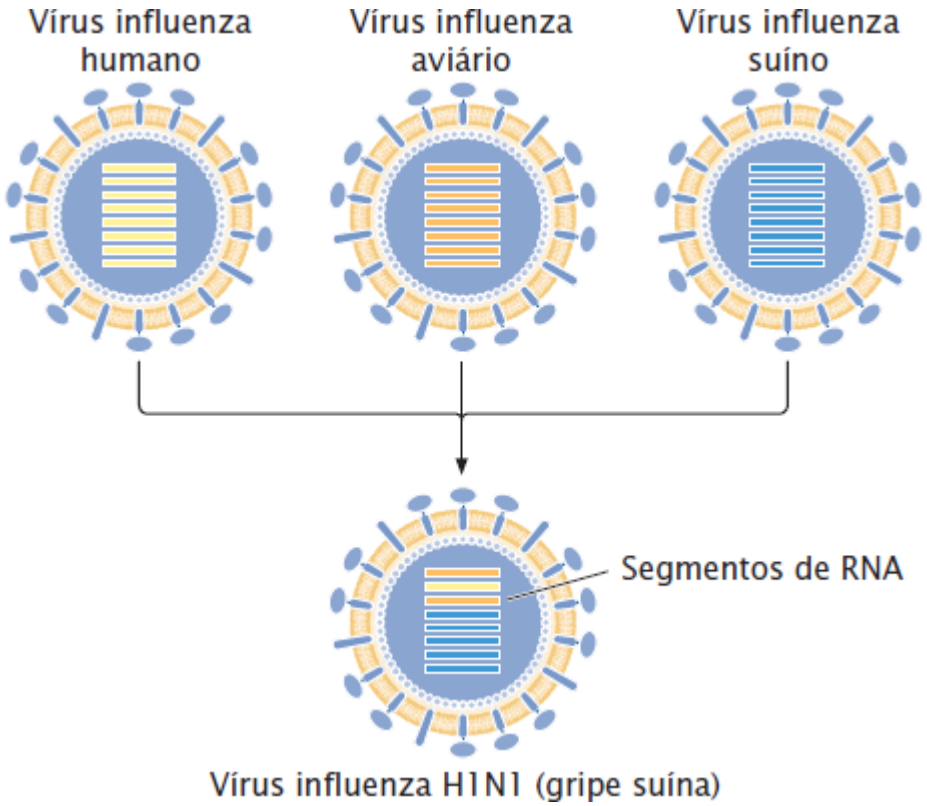


Figura 9.29 Novas cepas do vírus influenza são criadas pela recombinação do material genético de diferentes cepas. Um novo vírus H1N1 (gripe suína) que apareceu em 2009 tinha material genético dos vírus aviário, suíno e humano.

Resumo dos conceitos

- As bactérias e vírus são muito adequados para os estudos em genética: são pequenos, têm um pequeno genoma haploide, sofrem reprodução rápida e produzem muitos descendentes por reprodução assexuada
- O genoma bacteriano tem uma única molécula circular de DNA de fita dupla. Os plasmídios são pequenos pedaços de DNA bacteriano que podem se replicar independentemente do cromossomo
- O DNA pode ser transferido entre células bacterianas por meio da conjugação, transformação ou transdução
- A conjugação é a união de duas células bacterianas e a transferência de material genético entre elas. Ela é controlada por um episssomo chamado F. A taxa na qual os genes individuais são transferidos durante a conjugação indica a ordem dos genes e as distâncias entre eles no cromossomo bacteriano
- As bactérias captam DNA do ambiente por meio do processo de transformação. As frequências da cotransformação dos genes indicam as distâncias físicas entre os genes cromossômicos
- Já foram determinadas as sequências completas de DNA de muitas espécies de bactérias. Essas informações da sequência indicam que a transferência horizontal dos genes – o movimento de DNA entre as espécies – é comum nas bactérias
- Os vírus são estruturas replicantes com genomas DNA ou RNA que podem ser fita dupla ou única, linear ou circular
- Os genes bacterianos são incorporados em invólucros de fago e transferidos para outras bactérias por fagos pelo processo de transdução. As taxas de cotransdução podem ser usadas para mapear os genes bacterianos
- Os genes dos fagos podem ser mapeados ao infectar células bacterianas com duas cepas diferentes de fagos e contando o número de placas recombinantes produzidas pelos fagos descendentes
- Benzer mapeou várias mutações que ocorreram na região *rII* do fago T4. Os resultados de seus estudos de complementação demonstraram que a região *rII* tem duas unidades funcionais, as quais ele chamou de *cístrons*. Benzer mostrou que ocorre a recombinação intragênica
- Vários vírus têm genomas RNA. Os retrovírus codificam a transcriptase reversa, uma enzima usada para fazer uma cópia de DNA do genoma viral, que então se integra com o genoma do hospedeiro como um pró-vírus. O HIV é um retrovírus, e é o agente causador da AIDS
- A *influenza* é causada pelos vírus RNA que evoluem por meio de pequenas mudanças que ocorrem por mutação (variação

antigênica) e por mudanças mais expressivas que ocorrem por meio da recombinação do material genético a partir de diferentes cepas (mudança antigênica).

Termos importantes

Bactéria auxotrófica
Bactéria prototrófica
Célula competente
Colônia
Conjugação
Cotransduzido
Cotransformado
Episssomo
Fago temperado
Fago transdutor
Fago virulento
Fator F (fertilidade)
Integrase
Mapeamento intragênico
Meio de cultura completo
Meio de cultura mínimo
Mudança antigênica
Oncogene
Pilus (pl. *pilis*)
Placa
Plasmídio
Prófago
Pró-vírus
Retrovírus
Transcriptase reversa
Transdução
Transdução especializada
Transdução generalizada
Transdutante
Transferência horizontal de genes
Transformação
Transformante
Variação antigênica
Vírus

Respostas da Checagem dos conceitos

1. d.
2. b.
3. a.
4. *gal.*
5. *his* e *leu.*
6. b.
7. c.
8. b.
9. Transcriptase reversa.

Problema desenvolvido

Problema 1

O DNA de uma cepa de bactéria com o genótipo $a^+ b^+ c^+ d^+ e^+$ foi isolado e usado para transformar a cepa de bactérias $a^- b^- c^- d^- e^-$. As células transformadas foram testadas quanto à existência dos genes doados. Foram cotransformados os seguintes genes:

$$a^+ \text{ e } d^+b^+ \text{ e } e^+c^+ \text{ e } dc^+ \text{ e } e^+$$

Qual é a ordem dos genes *a, b, c, d* e *e* no cromossomo bacteriano?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Da ordem dos genes *a, b, c, d* e *e* no cromossomo bacteriano.

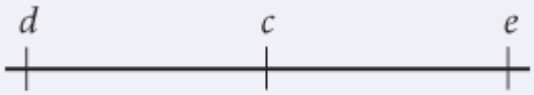
Quais informações são fornecidas para solucionar o problema?

- As células doadoras eram $a^+ b^+ c^+ d^+ e^+$ e as células receptoras eram $a^- b^- c^- d^- e^-$
- As combinações dos genes que foram cotransformados.

Para a solução deste problema, revise:
Transformação em bactérias, na Seção 9.2.

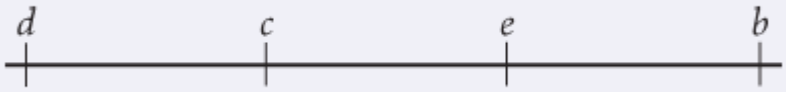
Etapas para a solução

Neste experimento de cotransformação o gene c^+ é cotransformado com os genes e^+ e d^+, mas os genes e^+ e d^+ não são cotransformados; portanto, o *locus c* está obrigatoriamente entre os *loci d* e *e*:

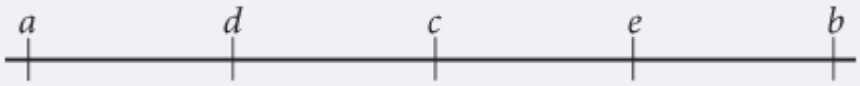


Lembrete: A taxa na qual os genes são cotransformados é inversamente proporcional à distância entre eles: os genes mais próximos são cotransformados com maior frequência, enquanto os que estão distantes são raramente cotransformados.

O gene e^+ também é cotransformado com o gene b^+; então os *loci e* e *b* devem estar próximos. O *locus b* poderia estar em qualquer um dos lados do *locus e*. Para determinar se o *locus b* está no mesmo lado de *e* como o *locus c*, procuramos ver se os genes b^+ e c^+ são cotransformados. Eles não são; então, o *locus b* está obrigatoriamente no lado oposto de *e* para *c*:



O gene a^+ é cotransformado com o gene d^+; ou seja, eles devem estar próximos. Se o *locus a* estiver no mesmo lado do *d* como o *locus c*, então os genes a^+ e c^+ seriam cotransformados. Como esses genes não têm cotransformação, o *locus a* está obrigatoriamente no lado oposto de *d*:



Problema 2

Considere três genes na *E. coli*: thr^+ (capacidade para sintetizar treonina), ara^+ (capacidade para metabolizar

arabinose) e *leu* (capacidade para sintetizar leucina). Esses três genes estão próximos no cromossomo da *E. coli*. Os fagos crescem em uma cepa thr^+ ara^+ leu^+ de bactéria (a cepa doadora). O lisado do fago é coletado e usado para infectar uma cepa de bactéria que é thr^- ara^- leu^-. As bactérias receptoras são testadas em um meio sem leucina. As bactérias que crescem e formam colônias nesse meio (transdutoras leu^+) são então replicadas em um meio sem treonina e um meio sem arabinose para ver quais são thr^+ e quais são ara^+.

Outro grupo de bactérias é testado em um meio sem treonina. As bactérias que crescem e formam colônias nesse meio (transdutantes thr^+) são então replicadas em um meio sem leucina e um meio sem arabinose para ver quais são ara^+ e quais são leu^+. Os resultados desses experimentos foram:

Marcador selecionado	Células com genes cotransduzidos (%)
leu^+	3 thr^+
	76 ara^+
thr^+	3 leu^+
	0 ara^+

Como os *loci* estão organizados no cromossomo?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Da ordem dos genes *thr*, *leu* e *ara* no cromossomo bacteriano.

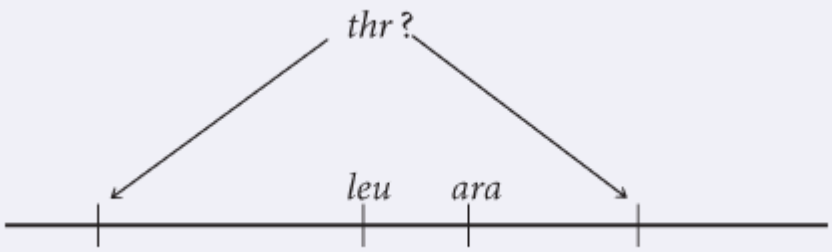
Quais informações são fornecidas para solucionar o problema?

- Os genes estão próximos no cromossomo da *E. coli*
- A cepa doadora é thr^+ ara^+ leu^+ e a cepa receptora é thr^- ara^- leu^-
- A porcentagem de células com genes cotransduzidos.

Para a solução deste problema, revise:
Transdução | *Como usar os fagos para mapear os genes bacterianos*, na Seção 9.3.

Etapas para a solução

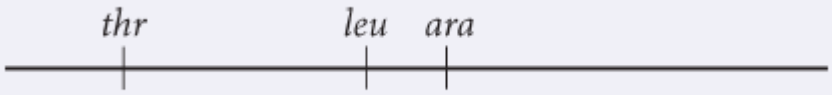
Observe que, quando selecionamos leu^+ (a metade superior do quadro), a maioria das células selecionadas também é ara^+. Esse achado indica que os genes *leu* e *ara* estão próximos, porque eles são cotransduzidos. Por outro lado, é raro thr^+ ser cotransduzido com leu^+, o que indica que *leu* e *thr* estão muito mais distantes. Com base nessas observações, sabemos que *leu* e *ara* estão mais próximos do que *leu* e *thr*, mas ainda não sabemos a ordem dos três genes – se *thr* está no mesmo lado de *ara* como *leu* ou se está no lado oposto, como indicado a seguir:



Dica: É mais provável que os genes próximos sejam cotransduzidos do que os genes que estão distantes.

Veja que, embora a frequência de cotransdução para *thr* e *leu* também seja 3%, não foram observados cotransdutores thr^+ ara^+. Esse achado indica que *thr* está mais próximo de *leu* do que *ara*, e, portanto, *thr*

deve estar à esquerda de *leu*, como indicado:



Dica: Podemos determinar a posição de *thr* com relação aos outros dois genes ao observar as frequências de transdução quando *thr*$^+$ é selecionado.

Questões de compreensão

Seção 9.1

1. Explique como as bactérias auxotróficas são isoladas.
2. Qual é a diferença entre um meio completo e um meio mínimo? Como os meios completo e mínimo nos quais foram adicionados um ou mais nutrientes (meios enriquecidos) são usados para isolar mutantes auxotróficos das bactérias?

Seção 9.2

3. Explique resumidamente as diferenças entre as células F^+, F^-, Hfr e F′.
4. Que tipos de cruzamentos são possíveis entre as células F^+, F^-, Hfr e F′? Quais são os desfechos desses cruzamentos? Qual é o papel do fator F na conjugação?
5. Explique como a conjugação interrompida, a transformação e a transdução podem ser usadas para mapear os genes bacterianos. Quais são as semelhanças e diferenças desses métodos?
6. O que é transferência horizontal de genes e como ela ocorre?

Seção 9.3

7. Liste algumas das características que tornam as bactérias e os vírus os organismos ideais para vários tipos de estudos genéticos.
8. Que tipos de genomas os vírus têm?
9. Descreva rapidamente as diferenças entre o ciclo lítico dos fagos virulentos e o ciclo lisogênico dos fagos temperados.
10. Explique resumidamente como os genes nos fagos são mapeados.
11. Qual é a diferença entre a transdução especializada e a transdução generalizada?
12. Explique brevemente o método usado por Benzer para determinar se ocorreram duas mutações diferentes no mesmo *locus*.
13. Descreva rapidamente a estrutura genética de um retrovírus típico.
14. Explique como um retrovírus, que tem um genoma RNA, é capaz de se integrar ao material genético de um hospedeiro com genoma DNA.
15. Quais são as origens evolutivas do HIV-1 e HIV-2?
16. A maioria dos seres humanos não é infectada com muita facilidade pela gripe aviária. Como, então, as sequências de DNA da gripe aviária são incorporadas ao DNA da gripe humana?

Questões e problemas aplicados

Introdução

*17. Suponha que você queira comparar as espécies de bactérias que existem em um rio poluído com as espécies que existem em um rio não poluído. Tradicionalmente, as bactérias foram identificadas ao fazê-las crescer no laboratório e comparar suas propriedades físicas e bioquímicas. Você sabe que não é possível cultivar a maioria das bactérias existentes nos rios. Como é possível identificar as espécies nos dois rios sem cultivá-las em laboratório?

Seção 9.2

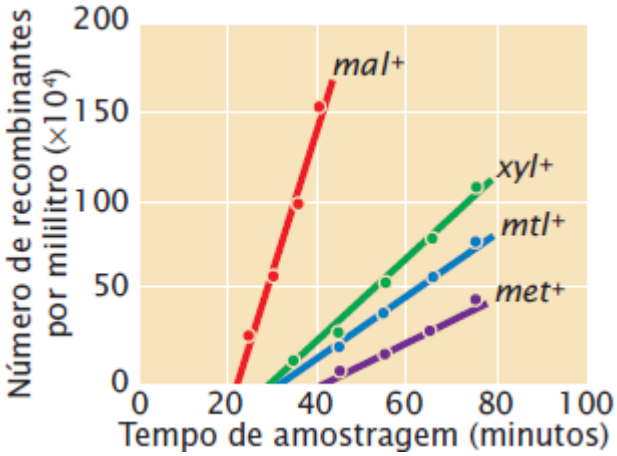
18. John Smith é um criador de porcos. Nos últimos cinco anos, ele adicionou vitaminas e pequenas doses de antibióticos na comida de seus porcos, e afirma que esses suplementos melhoram o crescimento dos animais. Entretanto, no último ano, vários porcos morreram de infecções por bactérias comuns, que não respondiam a grandes doses de antibióticos. Você

pode explicar essa taxa aumentada de mortalidade causada pela infecção nos porcos criados por Smith? Que conselho você daria a Smith para evitar esse problema no futuro?

19. Em casos raros, a conjugação das células Hfr e F^- produz duas células Hfr. Explique como ocorre esse evento.
20. Na Figura 9.7, o que as partes em vermelho e azul do cromossomo marcado pelo balão 6 representam?

ANÁLISE DE DADOS

21. Austim Taylor e Edward Adelberg isolaram novas cepas de células Hfr, e então as usaram para mapear vários genes na *E. coli* considerando a conjugação interrompida (A. L. Taylor and E. A. Adelberg. 1960. *Genetics* 45:1233-1243). Em um experimento, eles misturaram as células da cepa AB-312 Hfr, que eram *xyl*$^+$ *mtl*$^+$ *mal*$^+$ *met*$^+$ e sensíveis ao fago T6, com a cepa AB-531 F^-, que era *xyl*$^-$ *mtl*$^-$ *mal*$^-$ *met*$^-$ e resistente ao fago T6. Essas células sofreram conjugação. Em intervalos regulares, os pesquisadores removeram uma amostra e interromperam a conjugação ao exterminar as células Hfr com o fago T6. As células F^-, que eram resistentes ao fago T6, sobreviveram e então foram testadas quanto à existência dos genes transferidos da cepa Hfr. Os resultados desse experimento estão mostrados no gráfico a seguir. Com base nesses dados, indique a ordem dos genes *xyl*, *mtl*, *mal* e *met* no cromossomo bacteriano e indique as distâncias mínimas entre eles.



*22. Várias cepas Hfr com genótipo *m*$^+$ *n*$^+$ *o*$^+$ *p*$^+$ *q*$^+$ *r*$^+$ foram misturadas com uma cepa F^- com o genótipo *m*$^-$ *n*$^-$ *o*$^-$ *p*$^-$ *q*$^-$ *r*$^-$. A conjugação é interrompida em intervalos regulares, e a ordem de surgimento dos genes a partir da cepa Hfr é determinada nas células receptoras. A ordem da transferência de genes para cada cepa Hfr é:

Hfr5	*m*$^+$ *q*$^+$ *p*$^+$ *n*$^+$ *r*$^+$ *o*$^+$
Hfr4	*n*$^+$ *r*$^+$ *o*$^+$ *m*$^+$ *q*$^+$ *p*$^+$
Hfr1	*o*$^+$ *m*$^+$ *q*$^+$ *p*$^+$ *n*$^+$ *r*$^+$
Hfr9	*q*$^+$ *m*$^+$ *o*$^+$ *r*$^+$ *n*$^+$ *p*$^+$

Qual é a ordem dos genes no cromossomo bacteriano circular? Para cada cepa Hfr, indique a localização do fator F no cromossomo e sua polaridade.

*23. Foram realizados cruzamentos de três cepas diferentes de Hfr com amostras separadas de uma cepa F^-, e os seguintes dados de mapeamento foram obtidos a partir dos estudos da conjugação interrompida:

		Apresentação dos genes nas células F^-				
Hfr1:	Genes	*b*$^+$	*d*$^+$	*c*$^+$	*f*$^+$	*g*$^+$
	Tempo	3	5	16	27	59
Hfr2:	Genes	*e*$^+$	*f*$^+$	*c*$^+$	*d*$^+$	*b*$^+$
	Tempo	6	24	35	46	48
Hfr3:	Genes	*d*$^+$	*c*$^+$	*f*$^+$	*e*$^+$	*g*$^+$
	Tempo	4	15	26	44	58

Crie um mapa genético para esses genes, indicando sua ordem no cromossomo bacteriano e as distâncias entre eles.

24. Na Figura 9.14, qual gene no fator F será transferido por último na cepa Hfr5?

*25. O DNA de uma cepa de *Bacillus subtilis* com o genótipo *trp*$^+$ *tyr*$^+$ foi usado para transformar uma cepa receptora com o genótipo *trp*$^-$ *tyr*$^-$. Foram recuperados os seguintes números de células transformadas:

Genótipo	Número de células transformadas
trp$^+$ *tyr*$^-$	154
trp$^-$ *tyr*$^+$	312
trp$^+$ *tyr*$^+$	354

O que esses resultados sugerem sobre a ligação entre os genes *trp* e *tyr*?

26. O DNA de uma cepa de *Bacillus subtilis* com genótipo *a*$^+$ *b*$^+$ *c*$^+$ *d*$^+$ *e*$^+$ é usada para transformar uma cepa com genótipo *a*$^-$ *b*$^-$ *c*$^-$ *d*$^-$ *e*$^-$. Os pares de genes foram verificados quanto à cotransformação e foram obtidos os resultados a seguir:

Par de genes	Cotransformação	Par de genes	Cotransformação
a$^+$ e *b*$^+$	Não	*b*$^+$ e *d*$^+$	Não
a$^+$ e *c*$^+$	Não	*b*$^+$ e *e*$^+$	Sim
a$^+$ e *d*$^+$	Sim	*c*$^+$ e *d*$^+$	Não
a$^+$ e *e*$^+$	Sim	*c*$^+$ e *e*$^+$	Sim
b$^+$ e *c*$^+$	Sim	*d*$^+$ e *e*$^+$	Não

Com base nesses resultados, qual é a ordem dos genes no cromossomo bacteriano?

27. O DNA de uma cepa bacteriana *his*$^+$ *leu*$^+$ *lac*$^+$ é usado para transformar uma cepa *his*$^-$ *leu*$^-$ *lac*$^-$. Foram transformadas as seguintes porcentagens de células:

Cepa doadora	Cepa receptora	Genótipo das células transformadas	Porcentagem
his$^+$ *leu*$^+$ *lac*$^+$	*his*$^-$ *leu*$^-$ *lac*$^-$	*his*$^+$ *leu*$^+$ *lac*$^+$	0,02
		his$^+$ *leu*$^+$ *lac*$^-$	0,00
		his$^+$ *leu*$^-$ *lac*$^+$	2,00
		his$^+$ *leu*$^-$ *lac*$^-$	4,00
		his$^-$ *leu*$^+$ *lac*$^+$	0,10
		his$^-$ *leu*$^-$ *lac*$^+$	3,00
		his$^-$ *leu*$^+$ *lac*$^-$	1,50

a. Que conclusões você pode tirar sobre a ordem desses três genes no cromossomo?
b. Que genes estão mais próximos?

28. Rollim Hotchkiss e Julius Marmur estudaram a transformação na bactéria *Streptococcus pneumoniae* (R. D. Hotchkiss and J. Marmur. 1954. *Proceedings of the National Academy of Sciences* 40:55-60). Eles examinaram quatro mutações nessa bactéria: resistência à penicilina (*P*), resistência à estreptomicina (*S*), resistência à sulfanilamida (*F*) e a capacidade de produzir manitol (*M*). Os autores extraíram DNA de cepas das bactérias com diferentes combinações de diferentes mutações e usaram esse DNA para transformar células bacterianas do tipo selvagem (*P*$^+$ *S*$^+$ *F*$^+$ *M*$^+$). Os resultados a partir de seus experimentos de transformação são mostrados a seguir.

DNA da célula doadora	DNA da célula receptora	Transformantes	Porcentagem de todas as células
M S F	*M*$^+$ *S*$^+$ *F*$^+$	*M*$^+$ *S F*$^+$	4,0
		M$^+$ *S*$^+$ *F*	4,0
		M S$^+$ *F*$^+$	2,6
		M S F$^+$	0,41
		M$^+$ *S F*	0,22
		M S$^+$ *F*	0,0058
		M S F	0,0071

a. Hotchkiss e Marmur perceberam que a porcentagem de cotransformação era maior que a esperada de forma aleatória. Por exemplo, os resultados mostram que 2,6% das células foram transformadas em *M* e 4% foram transformadas em *S*. Se os traços *M* e *S* foram herdados de forma independente, a probabilidade esperada da cotransformação de *M* e *S* (*M S*) seria 0,026 × 0,04 = 0,001, ou 0,1%. Entretanto, eles observaram 0,41% de cotransformantes *M S*, quatro vezes

mais do que era esperado. Que fator é responsável pela frequência relativamente elevada de cotransformação dos traços que eles observaram?

b. Com base nos resultados, que conclusão você pode tirar sobre a ordem dos genes *M, S* e *F* no cromossomo bacteriano?

c. Por que a taxa de cotransformação para todos os três genes (*M S F*) é quase a mesma que a taxa de cotransformação para *M F* sozinho?

29. Durante o estudo dos efeitos do corte mecânico do DNA, Eugene Nester, A. T. Ganesan e Joshua Lederberg estudaram a transferência, por transformação, de DNA cortado de uma cepa do tipo selvagem de *Bacillus subtilis* (his_2^+ aro_3^+ try_2^+ aro_1^+ tyr_1^+ aro_2^+) para cepas de bactérias com uma série de mutações (E. W. Nester, A. T. Ganesan, and J. Lederberg. 1963. *Proceedings of the National Academy of Sciences* 49:61-68). Eles relataram as taxas de cotransformação entre his_2^+ e outros genes (expressas como taxa de cotransferência).

Genes	Taxa de cotransferência
his_2^+ e aro_3^+	0,015
his_2^+ e try_2^+	0,10
his_2^+ e aro_1^+	0,12
his_2^+ e tyr_1^+	0,23
his_2^+ e aro_2^+	0,05

Com base nesses dados, qual gene está mais distante de his_2^+? Qual gene está mais perto?

30. C. Anagnostopoulos e I. P. Crawford isolaram e estudaram várias mutações que afetavam diversas etapas da via bioquímica que leva ao triptofano na bactéria *Bacillus subtilis* (C. Anagnostopoulos and I. P. Crawford. 1961. *Proceedings of the National Academy of Sciences* 47:378-390). Sete das cepas que eles usaram no estudo estão listadas aqui, junto com a mutação encontrada em cada uma.

Cepa	Mutação
T3	T^-
168	I^-
168PT	I^-
TI	I^-
TII	I^-
T8	A^-
H25	H^-

Para mapear os genes responsáveis pela síntese do triptofano, eles fizeram vários experimentos de transformação nas cepas com diferentes mutações e determinaram a porcentagem de recombinantes entre as bactérias transformadas. Seus resultados foram:

Célula receptora	Célula doadora	Percentagem de recombinantes
T3	168PT	12,7
T3	T11	11,8
T3	T8	43,5
T3	H25	28,6
168	H25	44,9
TII	H25	41,4
TI	H25	31,3
T8	H25	67,4
H25	T3	19,0
H25	TII	26,3
H25	TI	13,4
H25	T8	45,0

Com base nessas frequências de recombinação de dois pontos, determine a ordem dos genes e as distâncias entre eles. Onde mais de um cruzamento foi feito para um par de genes, calcule a média das taxas de recombinação para os diferentes cruzamentos. Desenhe um mapa dos genes nos cromossomos.

Seção 9.3

*31. Duas mutações que afetam a morfologia da placa nos fagos (a^- e b^-) foram isoladas. Os fagos com ambas as mutações ($a^- b^-$) foram misturados com os do tipo selvagem ($a^+ b^+$) e adicionados a uma cultura de células bacterianas. Após infecção e lise, as amostras de lisado de fago foram coletadas e cultivadas em células bacterianas. Foram observados os seguintes números de placas:

Fenótipo da placa	Número
$a^+ b^+$	2.043
$a^+ b^-$	320
$a^- b^+$	357
$a^- b^-$	2.134

Qual é a frequência de recombinação entre os genes *a* e *b*?

*32. T. Miyake e M. Demerec examinaram mutações que precisam de prolina na bactéria *Salmonella typhimurium* (T. Miyake and M. Demerec. 1960. *Genetics* 45:755-762). Com base nos estudos de complementação, eles descobriram quatro auxótrofos da prolina: *proA*, *proB*, *proC* e *proD*. Para determinar se os *loci* de *proA*, *proB*, *proC* e *proD* estavam próximos no cromossomo bacteriano, eles fizeram um experimento de transdução. As cepas bacterianas que eram $proC^+$ e tinham mutações em *proA*, *proB* e *proD* foram usadas como doadoras. Elas foram infectadas com bacteriófagos, e os fagos descendentes infectaram bactérias receptoras com o genótipo $proC^- proA^+ proB^+ proD^+$. As bactérias foram plaqueadas em um meio seletivo para permitir que apenas as bactérias $proC^+$ crescessem. Então, os transdutantes $proC^+$ foram plaqueados nos meios seletivos para revelar seus genótipos nos outros três *loci pro*. Foram obtidos os seguintes resultados:

Genótipo da doadora	Genótipo do transdutante	Número
$proC^+ proA^- proB^+ proD^{+p}$	$proC^+ proA^+ proB^+ proD^+$	2.765
	$proC^+ proA^- proB^+ proD^+$	3
$proC^+ proA^+ proB^- proD^+$	$proC^+ proA^+ proB^+ proD^+$	1.838
	$proC^+ proA^+ proB^- proD^+$	2
$proC^+ proA^+ proB^+ proD^-$	$proC^+ proA^+ proB^+ proD^+$	1.166
	$proC^+ proA^+ proB^+ proD^-$	0

a. Por que não há genótipos $proC^-$ entre os transdutantes?
b. Quais genótipos representam transdutantes únicos e quais representam cotransdutantes?
c. Existem indícios de que *proA*, *proB* e *proD* estejam próximos de *proC*? Explique sua resposta.

*33. Uma geneticista isolou duas mutações em um bacteriófago. Uma mutação produz placas claras (*c*), e a outra produz placas pequenas (*m*). Os experimentos de mapeamento anteriores estabeleceram que a distância entre os genes responsáveis por essas duas mutações é 8 m.u. A geneticista misturou os fagos com o genótipo $c^+ m^+$ e genótipo $c^- m^-$ e usou a mistura para infectar bactérias. Ela coletou os fagos descendentes e cultivou uma amostra deles em bactérias plaqueadas. Foi observado um total de 1.000 placas. Quantos diferentes tipos de placas ($c^+ m^+$, $c^- m^-$, $c^+ m^-$, $c^- m^+$) ela espera observar?

34. A geneticista fez o experimento descrito no problema 33, mas desta vez ela misturou os fagos com os genótipos $c^+ m^-$ e $c^- m^+$. Quais resultados são esperados a partir deste cruzamento?

*35. Um geneticista isolou dois bacteriófagos mutantes *r* (r_{13} e r_2) que provocam lise rápida. Ele fez os cruzamentos a seguir e contou o número de placas listadas aqui:

Genótipo do fago parental	Descendentes	Número de placas
$h^+ r_{13}^- \times h^- r_{13}^+$	$h^+ r_{13}^+$	1
	$h^- r_{13}^+$	104
	$h^+ r_{13}^-$	110

	$h^- r_{13}^-$	2
Total		216
$h^+ r_2^- \times h^- r_2^+$	$h^+ r_2^+$	6
	$h^- r_2^+$	86
	$h^+ r_2^-$	81
	$h^- r_2^-$	7
Total		180

a. Calcule as frequências de recombinação entre r_2 e *h* e entre r_{13} e *h*.
b. Desenhe todos os possíveis mapas de ligação para esses três genes.

*36. Células da *E. coli* são simultaneamente infectadas com duas cepas do fago λ. Uma cepa tem uma faixa de hospedeiro mutante, é sensível à temperatura e produz placas claras (o genótipo é *h st c*); a outra cepa tem os alelos do tipo selvagem (o genótipo é $h^+ st^+ c^+$). Os fagos descendentes são coletados das células lisadas e plaqueados nas bactérias. Os números dos diferentes fagos descendentes estão listados a seguir:

Genótipo do fago descendente	Número de placas
$h^+ c^+ st^+$	321
h c st	338
h^+ *c st*	26
h c^+ st^+	30
h^+ *c* st^+	106
h c^+ *st*	110
h^+ c^+ *st*	5
h c st^+	6

a. Determine a ordem dos três genes no cromossomo do fago.
b. Determine as distâncias do mapa entre os genes.
c. Determine o coeficiente de coincidência e a interferência (ver *Capítulo 7*).

37. Uma cepa doadora de bactérias com os genes $a^+ b^+ c^+$ é infectada com fagos para mapear o cromossomo doador com a transdução generalizada. O lisado de fago a partir das células bacterianas é coletado e usado para infectar uma segunda cepa de bactérias $a^- b^- c^-$. As bactérias com o gene a^+ são selecionadas e a porcentagem de células com os genes b^+ e c^+ cotransduzido é registrada.

Doadora	Receptora	Gene selecionado	Células com gene cotransduzido (%)
$a^+ b^+ c^+$	$a^- b^- c^-$	a^+	25 b^+
		a^+	3 c^+

Qual gene, *b* ou *c*, está mais próximo do gene *a*? Explique.

38. Uma cepa de bactéria doadora com genótipo $leu^+ gal^- pro^+$ é infectada por fagos. O lisado de fagos a partir das células bacterianas é coletado e usado para infectar uma segunda cepa de bactérias $leu^- gal^+ pro^-$. A segunda cepa é selecionada para leu^+, e são obtidos os seguintes dados da cotransdução:

Doadora	Receptora	Gene selecionado	Células com gene cotransduzido (%)
$leu^+ gal^- pro^+$	$leu^- gal^+ pro^-$	leu^+	47 pro^+
		leu^+	26 gal^-

Quais genes estão mais próximos, *leu* e *gal* ou *leu* e *pro*?

39. Um geneticista isolou duas novas mutações, chamadas de rII_x e rII_y, a partir da região *rII* do bacteriófago T4. As células de *E. coli* B são simultaneamente infectadas com fagos que carreiam a mutação rII_x e com fagos que carreiam a mutação rII_y. Após as células se romperem, foram coletadas amostras de lisado de fagos. Uma amostra foi cultivada com células de *E. coli* K, e uma segunda amostra com as células de *E. coli* B. Existem 8.322 placas na B de *E. coli* e 3 placas na K de *E. coli*. Qual é a frequência de recombinação entre essas duas mutações?

*40. Um geneticista está trabalhando com um novo bacteriófago chamado fago Y3 que infecta a *E.coli*. Ele isolou oito fagos mutantes que não produzem placas quando crescem na cepa de células de *E. coli*. K Para determinar se essas mutações

ocorrem no mesmo gene funcional, ele infectou simultaneamente essas células com combinações pareadas dos mutantes e observou se as placas surgiam. Ele obteve os seguintes resultados. (Um sinal de mais significa que as placas foram formadas na célula de *E. coli* K e um sinal de menos indica que não foram formadas placas nessa cultura).

Mutante	1	2	3	4	5	6	7	8
1								
2	+							
3	+	+						
4	+	−	+					
5	−	+	+	+				
6	−	+	+	+	−			
7	+	−	+	−	+	+		
8	−	+	+	+	−	−	+	

a. Essas mutações pertencem a quantos genes funcionais (cístrons)?
b. Que mutações pertencem ao mesmo grupo funcional?

41. Para o vírus H1N1 da influenza mostrado na parte inferior da Figura 9.29, os vírus de qual organismo contribuíram a maior parte do RNA para o vírus H1N1?

Questões desafiadoras

Seção 9.2

42. Como projeto de verão, um estudante de microbiologia isola independentemente duas mutações na *E. coli* que são auxotróficas para glicina (*gly*$^-$). Ele queria saber se essas duas mutações estão na mesma unidade funcional. Demonstre um procedimento que o estudante poderia usar para determinar se essas duas mutações *gly*$^-$ ocorrem na mesma unidade funcional.
43. Um grupo de estudantes de genética misturou duas cepas auxotróficas de bactérias: uma é *leu*$^+$ *trp*$^+$ *his*$^-$ *met*$^-$ e a outra é *leu*$^-$ *trp*$^-$ *his*$^+$ *met*$^+$. Após misturar as duas cepas, eles plaquearam as bactérias em meio mínimo e observaram algumas colônias prototróficas (*leu*$^+$ *trp*$^+$ *his*$^+$ *met*$^+$). Os estudantes assumiram que houve transferência de genes entre as duas cepas. Como eles podem determinar se a transferência de genes ocorreu por conjugação, transdução ou transformação?

10

DNA | A Natureza Química do Gene

Viagens pelo Ártico e DNA antigo

A Groenlândia é a maior ilha do mundo, com 2.200.000 quilômetros quadrados, mas a maior parte da ilha está habitualmente soterrada sob muitos metros de gelo. Esse é um dos ambientes mais inóspitos da Terra. As temperaturas ao longo da costa sobem alguns graus acima do congelamento durante os dias de verão, mas caem para abaixo de zero durante boa parte do inverno. Com a luz do dia limitada (o sol se move acima do horizonte por apenas algumas horas nos dias de inverno), frio extremo e ventos alcançando a força de um ciclone, a Groenlândia é um ambiente perigosamente inóspito.

Apesar das condições severas, os moradores do Ártico ocupam continuamente a Groenlândia há quase 5.000 anos. Os primeiros habitantes foram os saqqaq, que ocuparam pequenos assentamentos na costa da Groenlândia por volta de 4.800 a 2.500 anos atrás. Eles viviam em pequenas tendas e caçavam mamíferos e aves marinhos. A origem desse povo foi um mistério por muito tempo. Afinal, eles descendiam de ameríndios, que migraram da Ásia para o Novo Mundo e depois se deslocaram para a Groenlândia? Ou descendiam do mesmo grupo que deu origem aos inuítes, que atualmente habitam regiões em torno do Círculo Polar Ártico? Ou talvez se originaram de outro grupo que migrou independentemente da Ásia para a Groenlândia após os ancestrais dos inuítes e dos nativos americanos terem chegado ao Novo Mundo?

O mistério da origem dos saqqaq foi solucionado em 2010, quando os geneticistas determinaram a sequência completa do DNA de um homem saqqaq de 4.000 anos – apelidado de Inuk –, cujos restos foram recuperados de um sítio arqueológico na costa oeste da Groenlândia. Os cientistas extraíram DNA de quatro tufos de cabelo encontrados na camada de terra congelada. Apesar da idade avançada da amostra, os cientistas conseguiram determinar a sequência completa do genoma de Inuk, que tinha mais de 3 bilhões de pares de base de DNA.

Ao comparar esse DNA com as sequências das populações conhecidas, os cientistas demonstraram que os saqqaq eram um parente próximo dos chukchis, um grupo atual de nômades nativos da Rússia. Esse achado indica que os saqqaq se originaram de caçadores que migraram da Sibéria oriental pelo Alasca e Canadá até a Groenlândia, chegando ao Novo Mundo independentemente de outros que deram origem aos ameríndios e aos inuítes. A análise posterior do DNA de Inuk revelou que ele tinha pele escura, olhos castanhos e tipo sanguíneo A+. Além disso, provavelmente estava ficando careca.

O DNA, com sua espiral de fita dupla, está entre as mais elegantes moléculas biológicas, mas a dupla-hélice não é apenas uma bela estrutura; ela também permite a notável estabilidade e permanência do DNA, como indicado pelo sequenciamento de um DNA de 4.000 anos. Em uma façanha ainda mais incrível, em 2009 os geneticistas sequenciaram o genoma completo do Neanderthal a partir do DNA extraído de ossos com 38.000 anos.

Este capítulo discute como o DNA foi identificado como a fonte de informação genética e como ele codifica as instruções genéticas. Começaremos considerando as exigências básicas do material genético e a história do estudo do DNA – como sua relação com os genes foi descoberta e sua estrutura, determinada. A história do DNA ilustra vários pontos importantes sobre a natureza da pesquisa científica. Assim como muitos avanços científicos importantes, a estrutura do DNA e seu papel como material genético não foram descobertos apenas por uma pessoa, mas foram gradualmente revelados durante o período de quase 100 anos, graças ao trabalho de vários investigadores. Nossa compreensão da relação entre o DNA e os genes foi aprimorada em 1953, quando James Watson e Francis Crick, ao analisarem os dados fornecidos por Rosalind Franklin e Maurice Wilkins, propuseram uma estrutura tridimensional para o DNA que esclareceu de forma brilhante seu papel na genética.

Examinaremos a estrutura do DNA após revisar as descobertas que levaram à nossa atual compreensão. A estrutura do DNA é importante por si só, mas o conceito-chave genético é a relação entre a estrutura e a função do DNA – como sua estrutura possibilitou que ele servisse de material genético.

10.1 O material genético tem várias características importantes

A vida é caracterizada por uma incrível diversidade, mas as instruções de codificação para todos os organismos vivos estão escritas na mesma linguagem genética – a dos ácidos nucleicos. É surpreendente, mas a ideia de que os genes são feitos de ácidos nucleicos não era muito aceita ainda depois de 1950. Esse ceticismo ocorria em parte por causa da falta de conhecimento sobre a estrutura do ácido desoxirribonucleico (DNA). Até a estrutura do DNA ser compreendida, não estava claro como ele podia armazenar e transmitir as informações genéticas. Mesmo antes de os ácidos nucleicos serem identificados como material genético, os biólogos reconheciam que, independentemente da natureza do material genético, ele tinha quatro características importantes.

1. **O material genético tem obrigatoriamente informações complexas.** Em primeiro lugar, o material genético tem de ser capaz de armazenar muita informação – instruções para os traços e as funções de um organismo.
2. **O material genético precisa se replicar de forma confiável.** Uma segunda característica necessária é a capacidade de o material genético ser copiado com exatidão. Cada organismo começa como uma única célula. Para produzir um organismo multicelular complexo como você, essa única célula precisa sofrer bilhões de divisões celulares. Em cada divisão celular, as instruções genéticas têm de ser transmitidas para as células descendentes com grande precisão. Quando os organismos se reproduzem e transferem os genes para seus descendentes, as instruções da codificação têm de ser copiadas com fidelidade.
3. **O material genético obrigatoriamente codifica o fenótipo.** O material genético (o genótipo) precisa ter a capacidade de ser expresso – para codificar os traços (o fenótipo). O produto de um gene é uma proteína ou uma molécula de RNA, então é crucial que exista um mecanismo para que as instruções genéticas no DNA sejam copiadas em RNAs e proteínas.
4. **O material genético precisa ter a capacidade de variar.** As informações genéticas têm de ser capazes de variar, porque diferentes espécies e até membros de uma espécie têm uma composição genética diferente.

Conceitos

O material genético tem de ser capaz de carrear numerosas informações, replicar-se de forma confiável, expressar suas instruções de codificação como fenótipos e ter a capacidade de variar.

Checagem dos conceitos 1

Por que a descoberta da estrutura do DNA foi tão importante para compreender a genética?

10.2 Todas as informações genéticas são codificadas na estrutura do DNA ou RNA

Embora nossa compreensão sobre como o DNA codifica a informação genética seja relativamente recente, o estudo da estrutura do DNA se desenrola há mais 100 anos (**Figura 10.1**).

Estudos iniciais sobre o DNA

Em 1868, Johann Friedrich Miescher se formou em Medicina na Suíça. Influenciado por um tio que acreditava que a chave para compreender as doenças estava na química dos tecidos, Miescher viajou para Tubingen, Alemanha, para estudar com Ernst Felix Hoppe-Seyler, um líder novato no emergente campo da bioquímica. Sob a orientação de Hoppe-Seyler, Miescher voltou sua atenção para a química do pus, uma substância de grande importância médica. O pus tem leucócitos com grandes núcleos, e Miescher desenvolveu um método para isolar esses núcleos. As pequenas quantidades de material nuclear que ele obteve foram insuficientes para uma análise química completa, mas Miescher estabeleceu que o material nuclear tinha uma substância nova levemente ácida e com elevado teor de fósforo. A esse material, composto de DNA e proteínas, Miescher deu o nome de *nucleína*. Posteriormente, a substância foi renomeada para *ácido nucleico* por um de seus alunos.

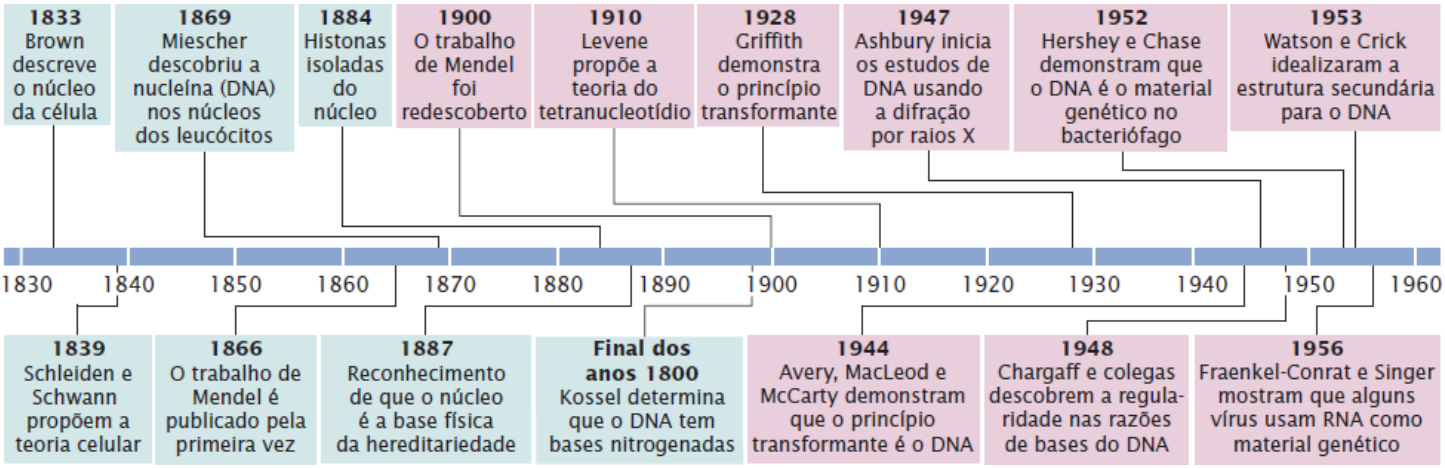
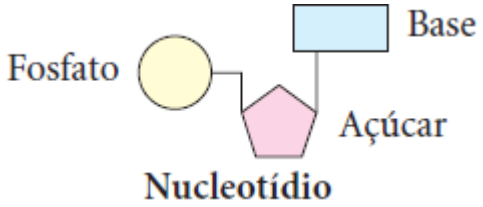


Figura 10.1 **Muitas pessoas contribuíram para nossa compreensão sobre a estrutura do DNA.**

Em 1887, vários pesquisadores concluíram, de forma independente, que a base física da hereditariedade está no núcleo. Demonstrou-se que a cromatina era composta de ácido nucleico e proteínas, mas não estava claro qual dessas substâncias era de fato a informação genética. No final do século 19, Albrecht Kossel fez um trabalho adicional sobre a química do DNA, e determinou que a estrutura tem quatro bases nitrogenadas: adenina, citosina, guanina e timina (abreviados como A, C, G e T).

No início do século 20, o Rockefeller Institute, na cidade de Nova York, tornou-se um centro para pesquisa sobre o ácido nucleico. Phoebus Aaron Levene ingressou no Rockefeller Institute em 1905 e passou os 40 anos seguintes estudando a química do DNA. Ele descobriu que o DNA tem um grande número de unidades repetidas, ligadas, chamadas **nucleotídios**; cada nucleotídio tem um açúcar, um fosfato e uma base.



Levene propôs, de forma incorreta, que o DNA tinha uma série de quatro unidades de nucleotídios, cada uma contendo todas as quatro bases – adenina, guanina, citosina e timina – em uma sequência fixa. Esse conceito, conhecido como a hipótese do tetranucleotídio, sugeria que a estrutura do DNA não é variável o suficiente para ser o material genético. A hipótese do tetranucleotídio contribuiu com a ideia de que a proteína é o material genético porque, com seus 20 diferentes aminoácidos, a estrutura da proteína poderia ser muito variável.

À medida que estudos adicionais sobre a química do DNA eram concluídos, nas décadas de 1940 e 1950, a noção do DNA como uma molécula simples e invariável começou a mudar. Erwin Chargaff e seus colegas mediram a quantidade das quatro bases no DNA a partir de vários organismos e descobriram que o DNA de diferentes organismos apresenta uma composição de bases muito diferente. Esse achado derrubou a hipótese do tetranucleotídio. Eles descobriram que, dentro de cada espécie, há uma regularidade nas razões das bases: a quantidade de adenina é sempre igual à quantidade de timina (A = T), e a quantidade de guanina é sempre igual à quantidade de citosina (G = C; **Quadro 10.1**). Esses achados ficaram conhecidos como as **regras de Chargaff**.

Conceitos

Detalhes da estrutura de DNA foram descobertos por vários cientistas. Primeiro, o DNA foi interpretado como sendo uma estrutura muito regular para carrear a informação genética, mas, nos anos de 1940, demonstrou-se que o DNA de diferentes organismos tinha composição de base variável.

Checagem dos conceitos 2

Qual foi a contribuição de Levene para nossa compreensão sobre a estrutura do DNA?

a. Determinou que o núcleo tem DNA.
b. Determinou que o DNA tem quatro bases nitrogenadas.
c. Determinou que o DNA é composto por nucleotídios.
d. Determinou que as bases dos nucleotídios do DNA existem em razões regulares.

DNA como fonte de informação genética

Enquanto os químicos descobriam a estrutura do DNA, os biólogos tentavam identificar a fonte da informação genética. Mendel identificou as regras básicas da hereditariedade em 1866, mas ele não tinha ideia da estrutura física da informação da hereditariedade. No início do século 20, os biólogos concluíram que os genes residiam nos cromossomos, conhecidos por terem DNA e proteína. Dois conjuntos de experimentos, um feito em bactérias e outro em vírus, forneceram evidências importantes de que o DNA, em vez da proteína, era o material genético.

Descoberta do princípio transformante. A primeira dica de que o DNA era o portador da informação da hereditariedade veio com a demonstração de que ele era o responsável por um fenômeno chamado *transformação*. Esse fenômeno foi observado pela primeira vez em 1928 por Fred Griffith, um médico inglês cujo interesse especial era a bactéria que causa pneumonia: *Streptococcus pneumoniae*. Ele conseguiu isolar diferentes cepas de *S. pneumoniae* (tipos I, II, III e assim em diante). Nas formas virulentas (causadores de doenças) de uma cepa, cada bactéria era circundada por um revestimento de polissacarídio, que torna a aparência da colônia bacteriana lisa quando cresce em uma placa de ágar; essas formas eram chamadas de S, do inglês *smooth*. Griffith descobriu que essas formas virulentas sofriam mutação para formas não virulentas, sem revestimento de polissacarídio, e produziam uma colônia rugosa; essas formas eram chamadas de R, para rugosa.

Ele observou que pequenas quantidades das bactérias do tipo IIIS vivas injetadas nos camundongos causavam pneumonia e os animais morriam; quando examinou os camundongos mortos, Griffith encontrou grandes quantidades da bactéria do tipo IIIS em seu sangue (**Figura 10.2 A**). Quando injetava as bactérias do tipo IIR nos camundongos, eles não morriam e nenhuma bactéria era recuperada de seu sangue (**Figura 10.2 B**). Sabendo que a fervura matava todas as bactérias e destruía sua virulência, Griffith injetou nos camundongos grandes inóculos de bactérias do tipo IIIS mortas pelo calor. Os camundongos sobreviveram e as bactérias do tipo IIIS não foram isoladas de seu sangue (**Figura 10.2 C**).

Quadro 10.1 Composição das bases (porcentagem)* de DNA de diferentes fontes e as razões das bases.

Fonte do DNA	A	T	G	C	Razão A/T	Razão G/C	Razão (A + G)/(T + C)
E. coli	26,0	23,9	24,9	25,2	1,09	0,99	1,04
Levedura	31,3	32,9	18,7	17,1	0,95	1,09	1,00
Ouriço-do-mar	32,8	32,1	17,7	18,4	1,02	0,96	1,00
Rato	28,6	28,4	21,4	21,5	1,01	1,00	1,00
Humano	30,3	30,3	19,5	19,9	1,00	0,98	0,99

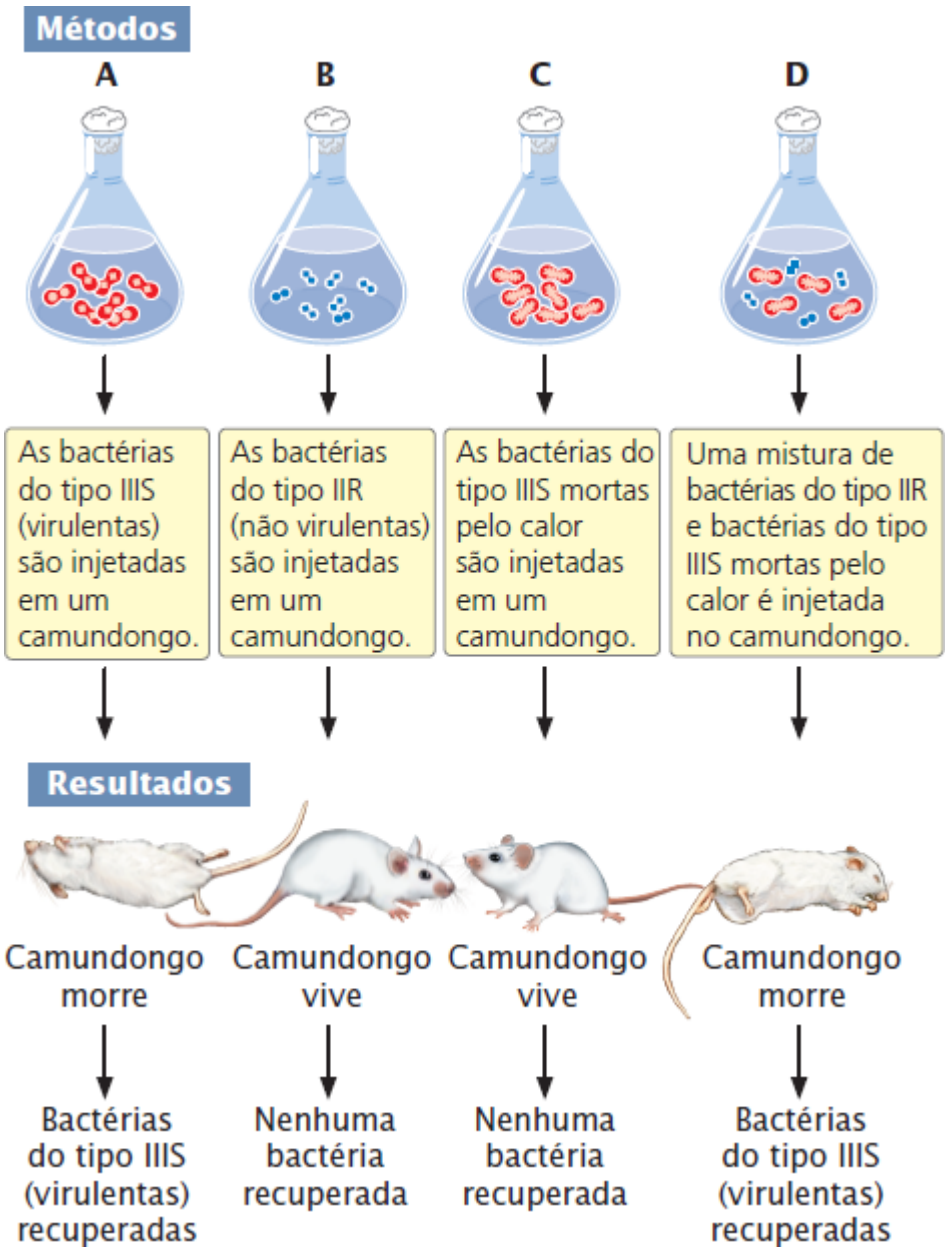
*Porcentagem em moles de constituintes nitrogenados por 100 g de átomos de fosfato no hidrolisado corrigido para 100% de recuperação. De E. Chargaff and J. Davidson (eds). The Nucleic Acids, Vol 1. (New York: Academic Press, 1955).

Os resultados desses experimentos não eram incomuns. Entretanto, Griffith teve uma surpresa quando infectou seus camundongos com uma pequena quantidade de bactérias do tipo IIR vivas e com uma grande quantidade de bactérias do tipo IIIS mortas pelo calor. Como as bactérias dos tipos IIR e IIIS mortas pelo calor não eram virulentas, ele esperava que esses camundongos vivessem. De maneira surpreendente, 5 dias após as injeções os camundongos tiveram pneumonia e morreram (**Figura 10.2 D**). Quando Griffith examinou o sangue dos corações desses camundongos, observou as bactérias do tipo IIIS vivas. Além disso, essas bactérias mantinham suas características IIIS por várias gerações; então, a infectividade era herdada.

Griffith considerou todas as interpretações possíveis de seus resultados. Primeiro, ele poderia não ter esterilizado suficientemente as bactérias do tipo IIIS e algumas bactérias vivas ficaram na cultura. Qualquer bactéria viva injetada nos camundongos teria se multiplicado e causado pneumonia. Ele sabia que essa possibilidade era improvável, porque havia usado apenas bactérias do tipo IIIS mortas pelo calor no experimento-controle, e elas não produziram pneumonia nos camundongos.

Experimento

Pergunta: É possível que um extrato de células bacterianas mortas transforme geneticamente células vivas?



Conclusão: Uma substância nas bactérias virulentas mortas pelo calor transformou geneticamente as bactérias do tipo IIR em bactérias do tipo IIIS virulentas e vivas.

Figura 10.2 Os experimentos de Griffith demonstraram a transformação nas bactérias.

Uma segunda interpretação era que as bactérias do tipo IIR vivas haviam sofrido mutação para a forma S virulenta. Essa mutação causaria pneumonia nos camundongos, mas produziria as bactérias do tipo IIS, não o tipo IIIS que Griffith encontrou nos camundongos mortos. Como as bactérias dos tipos II e III diferem em vários traços, seriam necessárias muitas mutações para que a bactéria do tipo II sofresse mutação para o tipo III, e a chance de que todas as mutações ocorressem ao mesmo tempo era extremamente baixa.

Griffith concluiu que as bactérias do tipo IIR tinham sido transformadas de alguma forma, adquirindo a virulência genética das IIIS mortas. Essa transformação produziu uma mudança genética permanente nas bactérias. Embora Griffith não compreendesse a natureza dessa transformação, ele desenvolveu uma teoria de que alguma substância no revestimento de polissacarídio da bactéria morta fosse a responsável. Ele chamou essa substância de **princípio transformante**. Resolva o Problema 19

Identificação do princípio transformante. Oswald Avery (ver Figura 10.1) era microbiologista no Rockefeller Institute quando Griffith publicou seu relatório. No primeiro momento, Avery ficou cético, mas após outros microbiologistas conseguirem repetir os experimentos de Griffith com outras bactérias, ele se propôs a identificar a natureza da substância

transformante.

Após 10 anos de pesquisa, Avery, Colin MacLeod e Maclyn McCarty conseguiram isolar e purificar parcialmente a substância transformante. Eles mostraram que ela apresentava uma composição química muito semelhante à do DNA e muito diferente das proteínas. Enzimas como tripsina e quimiotripsina, conhecidas por degradarem proteínas, não tinham efeito sobre essa substância. A ribonuclease, uma enzima que destrói RNA, também não tinha efeito. Entretanto, as enzimas capazes de destruir o DNA eliminavam a atividade biológica da substância transformante (**Figura 10.3**). Avery, MacLeod e McCarty mostraram que a substância transformante precipitava quase na mesma taxa que o DNA purificado e que absorvia luz ultravioleta nos mesmos comprimentos de onda que o DNA. Esses resultados, publicados em 1944, forneceram a prova definitiva de que o princípio transformante – e portanto, a informação genética – está no DNA. Entretanto, é muito raro novas teorias na ciência serem aceitas com base em um único experimento, e muitos biólogos ainda preferiam a hipótese de que o material genético fosse proteína.

Conceitos

O processo de transformação indica que alguma substância – o princípio transformante – é capaz de alterar geneticamente as bactérias. Avery, MacLeod e McCarty demonstraram que o princípio transformante é o DNA, fornecendo a primeira evidência de que o DNA é o material genético.

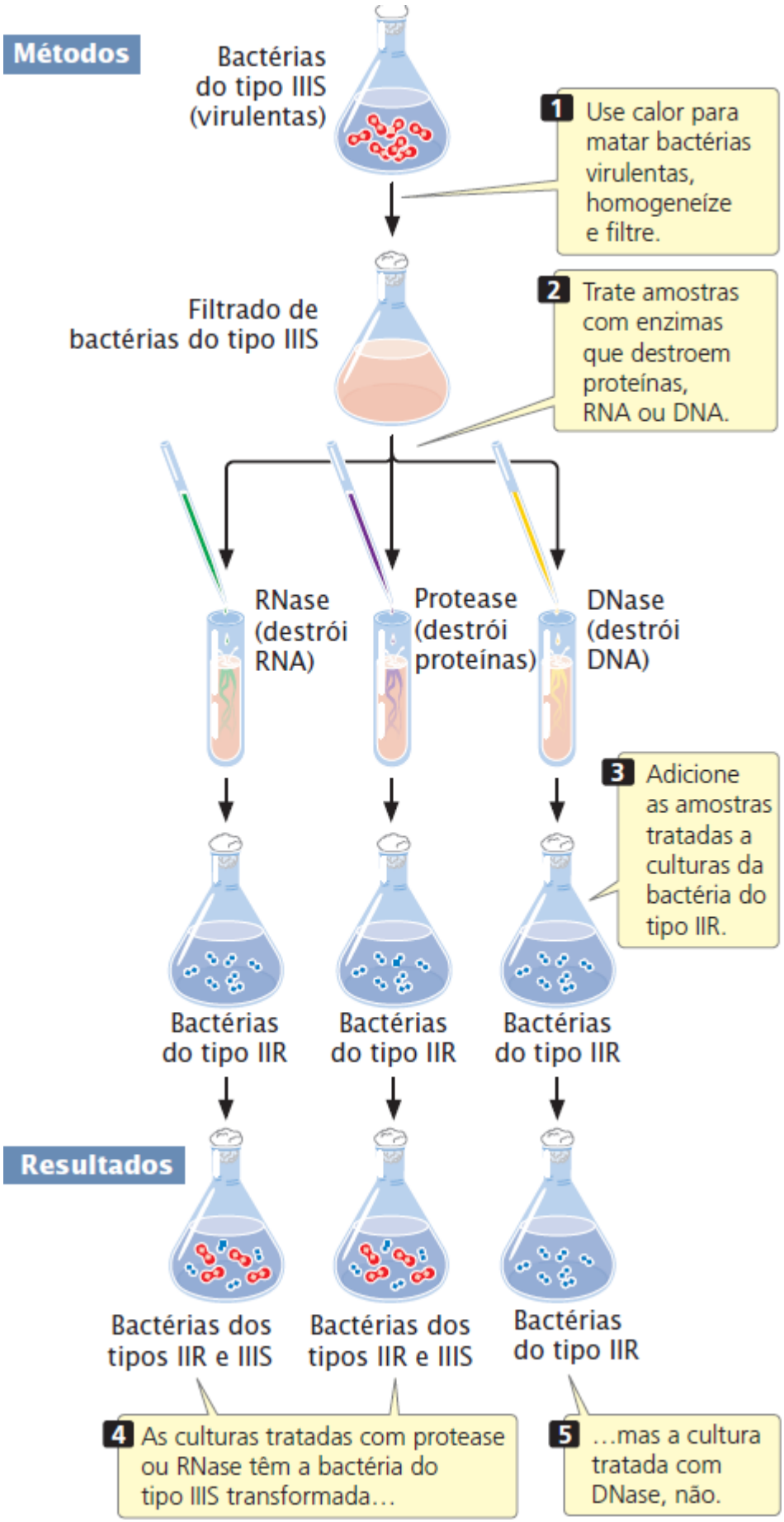
Checagem dos conceitos 3

Se Avery, MacLeod e McCarty descobrissem que as amostras de bactérias mortas com calor tratadas com RNase e DNase transformavam as bactérias, mas as amostras tratadas com protease não são transformadas, a que conclusões eles teriam chegado?

a. A protease é responsável pela transformação.
b. RNA e DNA são os materiais genéticos.
c. A proteína é o material genético.
d. RNase e DNase são necessárias para a transformação.

Experimento

Pergunta: Qual é a natureza química da substância transformante?

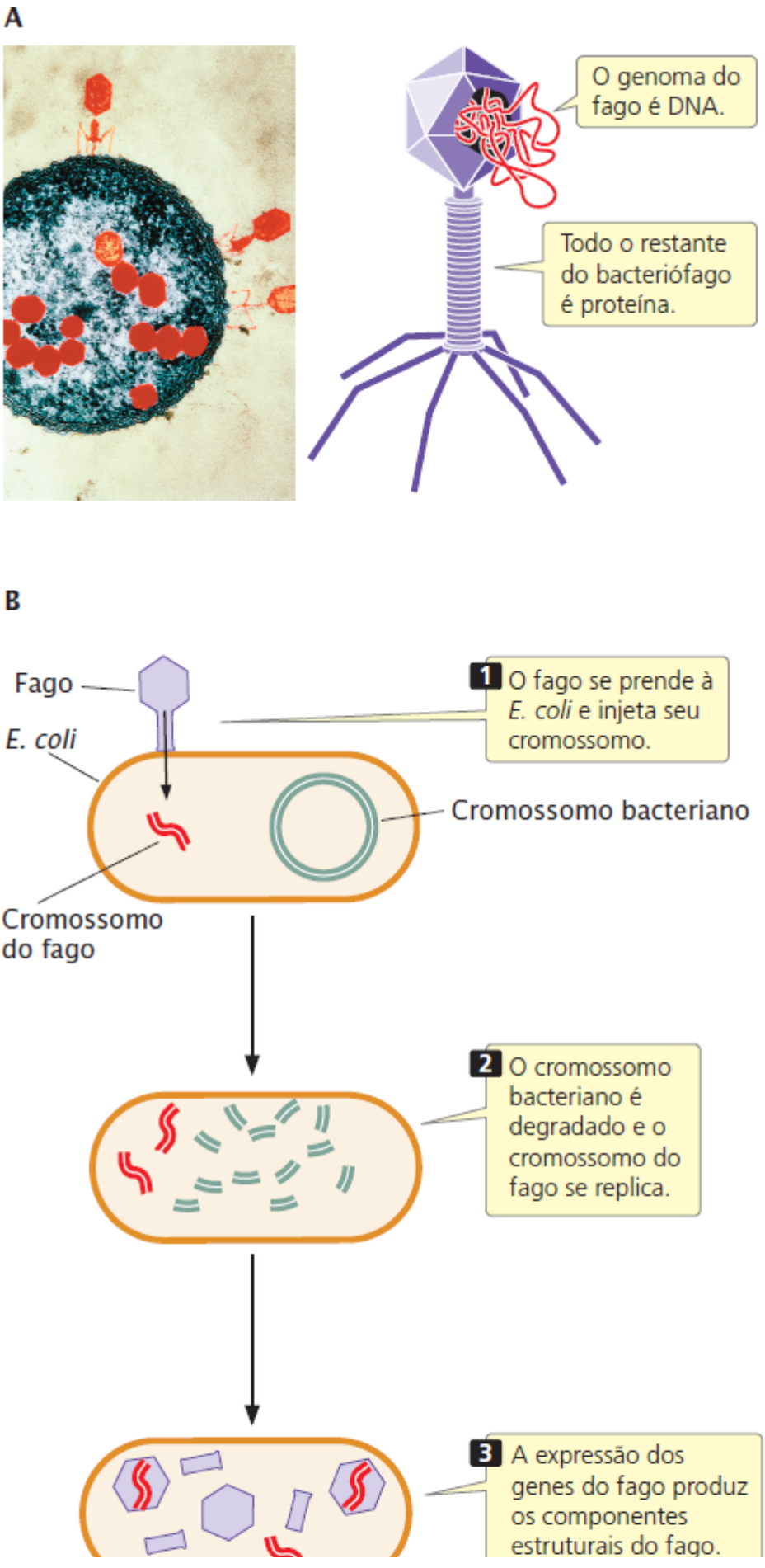


Conclusão: Como apenas a DNase destruiu a substância transformante, essa substância é o DNA.

Figura 10.3 O experimento de Avery, MacLeod e McCarty revelou a natureza do princípio transformante.

Experimento de Hershey-Chase. Uma segunda evidência de que o DNA era o material genético surgiu de um estudo do

vírus T2 conduzido por Alfred Hershey e Martha Chase. O vírus T2 é um *bacteriófago* (fago) que infecta a bactéria *Escherichia coli* (**Figura 10.4 A**). Como afirmado no Capítulo 9, um fago se reproduz ao se prender na parede externa de uma célula bacteriana e injetar seu DNA na célula, onde se replica e direciona a célula para sintetizar as proteínas do fago. O DNA do fago é encapsulado pelas proteínas, produzindo fagos descendentes que rompem a célula e escapam (**Figura 10.4 B**).



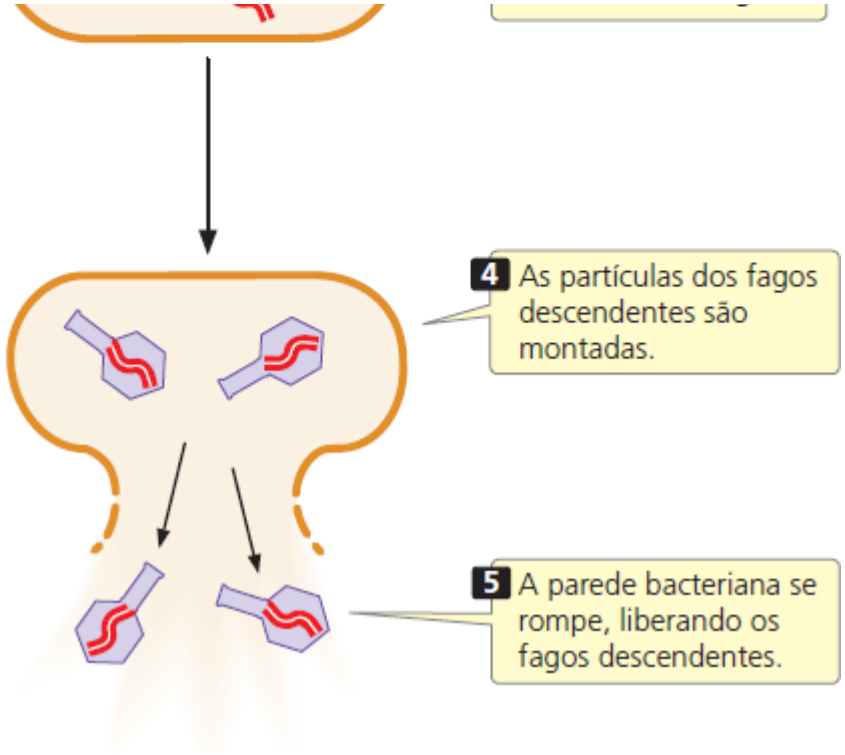


Figura 10.4 **T2 é um bacteriófago que infecta *E. coli.* A.** Fago T2. **B.** Seu ciclo de vida. (Micrografia: © Lee D. Simon/Photo researchers.)

No período do estudo de Hershey-Chase (o artigo deles foi publicado em 1952), os biólogos não compreendiam exatamente como os fagos se reproduziam. O que eles sabiam era que o fago T2 é aproximadamente 50% proteína e 50% DNA, que ele infecta uma célula ao se prender à parede celular e que os fagos descendentes eram produzidos dentro da célula. Como os descendentes carreiam os mesmos traços que o fago infectante, o material genético deste tem de ser transmitido para os descendentes, mas os pesquisadores não sabiam como essa transmissão genética ocorria.

Hershey e Chase desenvolveram vários experimentos para determinar se a *proteína* ou o DNA do fago é transmitido na reprodução. Para acompanhar o destino da proteína e do DNA, eles usaram formas radioativas, ou **isótopos**, de fósforo e enxofre. Um isótopo radioativo pode ser usado como marcador para identificar a localização de uma molécula específica porque qualquer molécula com o isótopo será radioativa e facilmente detectada. O DNA contém fósforo, mas não enxofre, então Hershey e Chase usaram ^{32}P para seguir o DNA do fago durante sua reprodução. A proteína contém enxofre, mas não fósforo; então foi usado o ^{35}S para seguir a proteína.

Os pesquisadores cresceram um lote de *E. coli* em um meio com ^{32}P e infectaram as bactérias com o fago T2, de modo que todos os novos fagos teriam um DNA marcado com ^{32}P (**Figura 10.5**). Eles cresceram um segundo lote de *E. coli* em um meio com ^{35}S e infectaram as bactérias com o fago T2, de modo que todos os novos fagos teriam proteína marcada com ^{35}S. Hershey e Chase infectaram lotes separados de *E. coli* não marcada com os fagos marcados com ^{35}S e ^{32}P. Aguardado um período para os fagos infectarem as células, eles colocaram as células de *E. coli* em um liquidificador e arrancaram os revestimentos de proteínas novos e vazios das paredes celulares; separaram os revestimentos de proteína e cultivaram as células bacterianas infectadas.

Quando fagos marcados com ^{35}S infectaram as bactérias, a maior parte da radioatividade foi detectada nos revestimentos de proteína e pouca radioatividade foi detectada nas células. Além disso, com os novos fagos emergindo da célula, eles não tinham quase nenhum ^{35}S (ver Figura 10.5). Esse resultado indicava que o componente proteico de um fago não entra na célula nem é transmitido para os fagos descendentes.

Por outro lado, quando Hershey e Chase infectaram as bactérias com fagos marcados com ^{32}P e removeram os revestimentos de proteína, as bactérias estavam radioativas. O mais impressionante: após as células se romperem e os novos fagos descendentes surgirem, muitos deses fagos emitiam radioatividade a partir do ^{32}P, demonstrando que o DNA dos fagos infectantes foi transmitido para os descendentes (ver Figura 10.5). Os resultados confirmam que o DNA, e não a proteína, é o material genético dos fagos. Resolva o Problema 24

Conceitos

Usando isótopos radioativos, Hershey e Chasey rastrearam o movimento do DNA e da proteína durante a infecção por fago. Eles demonstraram que o DNA – e não a proteína – entra na célula bacteriana durante a

reprodução do fago e que apenas o DNA é transmitido para os fagos descendentes.

Checagem dos conceitos 4

Hershey e Chase poderiam ter usado um isótopo radioativo do carbono em vez do ^{32}P? Por quê?

Descoberta da estrutura tridimensional do DNA por Watson e Crick

Os experimentos sobre a natureza do material genético definiram o estágio para um dos mais importantes avanços na história da biologia – a descoberta da estrutura tridimensional do DNA por James Watson e Francis Crick, em 1953.

Antes da descoberta de Watson e Crick, boa parte da química básica do DNA já havia sido determinada por Miescher, Kossel, Levene, Chargaff e outros, que estabeleceram que o DNA consistia em nucleotídios e que cada nucleotídio continha um açúcar, uma base e um grupo fosfato. Entretanto, não estava muito claro como os nucleotídios se encaixavam na estrutura tridimensional da molécula.

Em 1947, William Astbury começou a estudar a estrutura tridimensional do DNA ao usar uma técnica chamada **difração por raios X** (**Figura 10.6**), na qual os raios X que irradiam uma molécula são refletidos em padrões específicos que revelam aspectos da estrutura da molécula. Entretanto, suas imagens de difração não forneciam uma resolução adequada para revelar a estrutura. Um grupo de pesquisa no King's College em Londres, liderado por Maurice Wilkins, também usou a difração por raios X para estudar o DNA. Trabalhando no laboratório de Wilkins, Rosalind Franklin obteve imagens visivelmente melhores da molécula. Entretanto, o progresso de Wilkins e Franklin no desenvolvimento de uma estrutura completa da molécula foi prejudicado pela desavença pessoal entre eles.

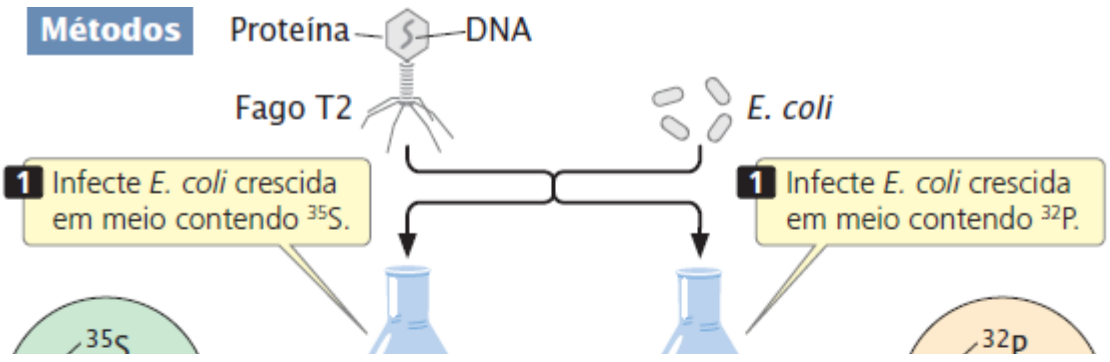
Watson e Crick investigavam a estrutura do DNA – não coletando novos dados, mas usando todas as informações disponíveis sobre a química do DNA – para construir modelos moleculares (**Figura 10.7 A**). Eles usaram imagens excelentes de difração por raios X tiradas por Rosalind Franklin (**Figura 10.7 B**), e ao aplicar as leis da química estrutural conseguiram limitar o número de estruturas possíveis que o DNA poderia assumir. Eles testaram várias estruturas ao construir modelos feitos de fios e placas de metal. Com seus modelos, foram capazes de ver se uma estrutura era compatível com os princípios da química e com as imagens de raios X.

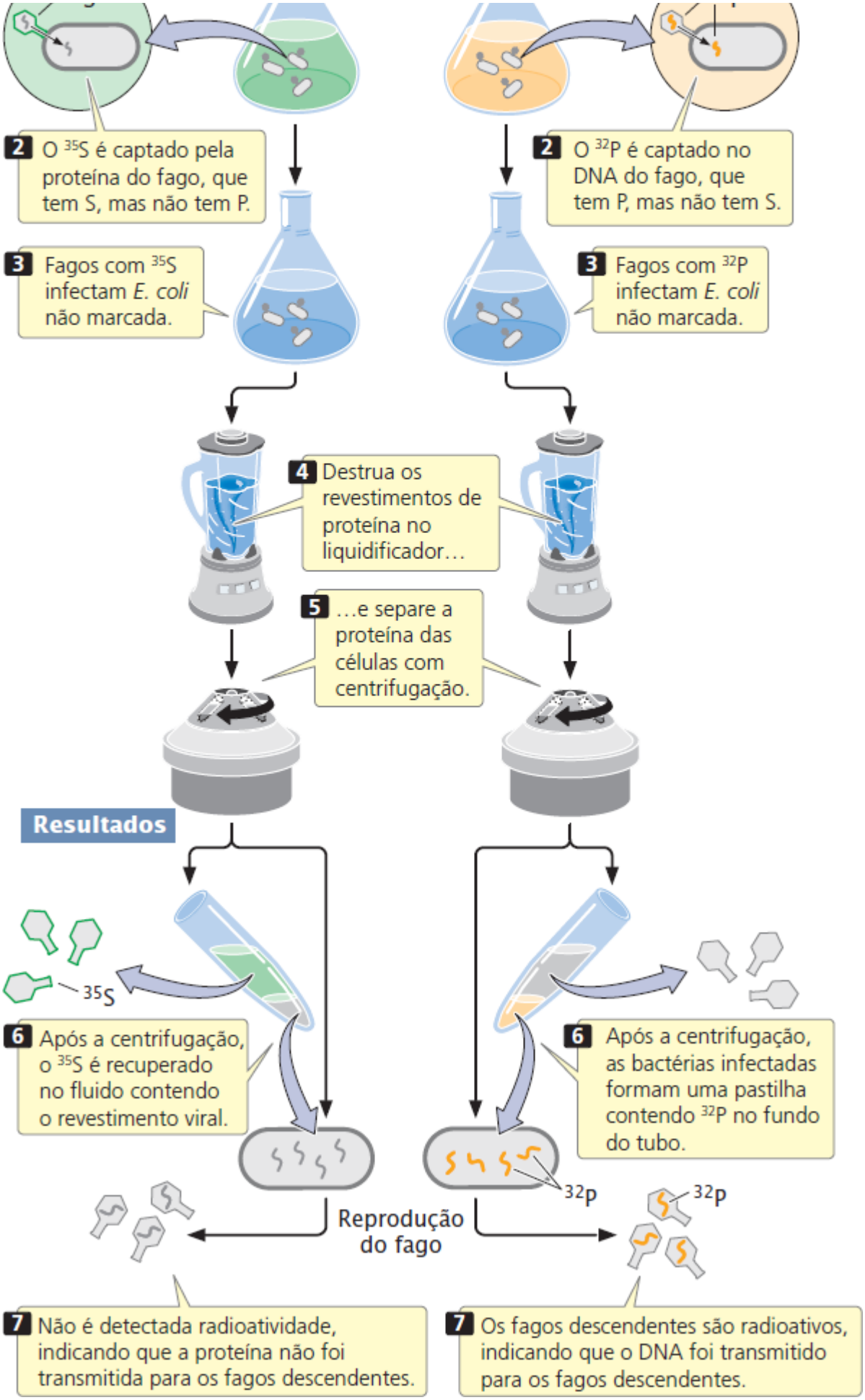
A chave para solucionar a estrutura veio quando Watson identificou que uma base adenina poderia se ligar com uma base timina e que uma base de guanina poderia se ligar com uma base citosina; esses pareamentos eram responsáveis pelas razões de base que Chargaff havia descoberto anteriormente. O modelo desenvolvido por Watson e Crick mostrou que o DNA era composto por duas cadeias de nucleotídios que estavam em direções opostas (são antiparalelas) e que se enrolavam uma sobre a outra para formar uma hélice girando para a direita, com os açúcares e fosfatos no lado externo e as bases no lado interno. Eles identificaram que a estrutura de fita dupla do DNA com seu pareamento de bases específico oferecia um meio elegante pelo qual o DNA pode ser replicado. Watson e Crick publicaram uma descrição fascinante do seu modelo na revista *Nature*, em 1953. Ao mesmo tempo, Wilkins e Franklin publicaram, cada um, seus dados sobre a difração por raios X, que demonstravam experimentalmente a hipótese de que o DNA tinha uma estrutura helicoidal.

Muitos dizem que a solução da estrutura do DNA é a descoberta da biologia mais importante do século 20. Por sua descoberta, Watson e Crick, com Maurice Wilkins, receberam o prêmio Nobel em 1962. Rosalind Franklin morreu de câncer em 1958 e não poderia ser considerada uma candidata ao prêmio compartilhado, mas muitos acadêmicos e historiadores acreditam que ela deveria receber o mesmo crédito por solucionar a estrutura do DNA.

Experimento

Pergunta: Que parte do fago – seu DNA ou sua proteína – serve como material genético e é transmitido para os descendentes?





Conclusão: O DNA, e não a proteína, é o material genético nos bacteriófagos.

Figura 10.5 Hershey e Chase demonstraram que o DNA carreia as informações genéticas nos bacteriófagos.

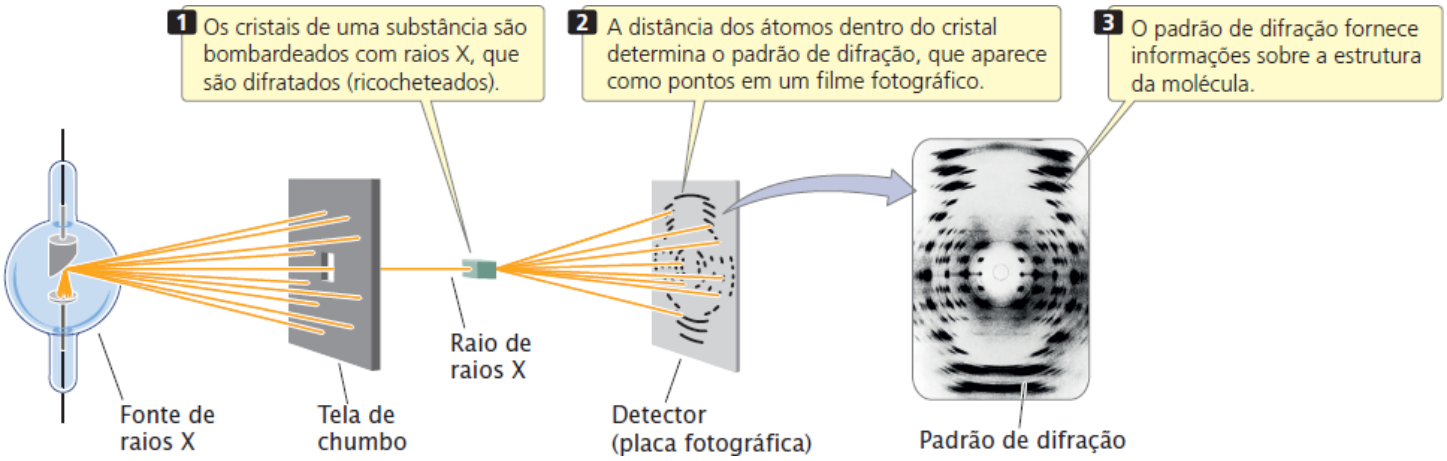


Figura 10.6 A difração por raios X fornece informações sobre as estruturas das moléculas. (Science Source.)

Figura 10.7 James Watson e Francis Crick (A) desenvolveram um modelo tridimensional da estrutura do DNA com base, em parte, nas fotografias da difração por raios X tiradas por Rosalind Franklin (B). (*Parte A*: A. Barrington Brown/Science Photo Library/Photo Researchers. *Parte B*: Science Source/Photo Researchers.)

Após a descoberta da estrutura do DNA, boa parte da pesquisa se concentra em como a informação genética é codificada na sequência de bases e como essa informação é copiada e expressa. Ainda hoje, detalhes da estrutura e da função do DNA são tema de pesquisa.

Conceitos

Ao coletar as informações sobre a química do DNA e construir modelos moleculares, Watson e Crick descobriram a estrutura tridimensional da molécula de DNA.

Checagem dos conceitos 5

O que Watson e Crick usaram para ajudá-los a descobrir a estrutura do DNA?

a. Difração por raios X.
b. Leis da estrutura química.
c. Modelos de DNA.
d. Todas as opções anteriores.

RNA como material genético

Na maioria dos organismos, o DNA carreia a informação genética. Entretanto, alguns vírus usam RNA em vez de DNA como

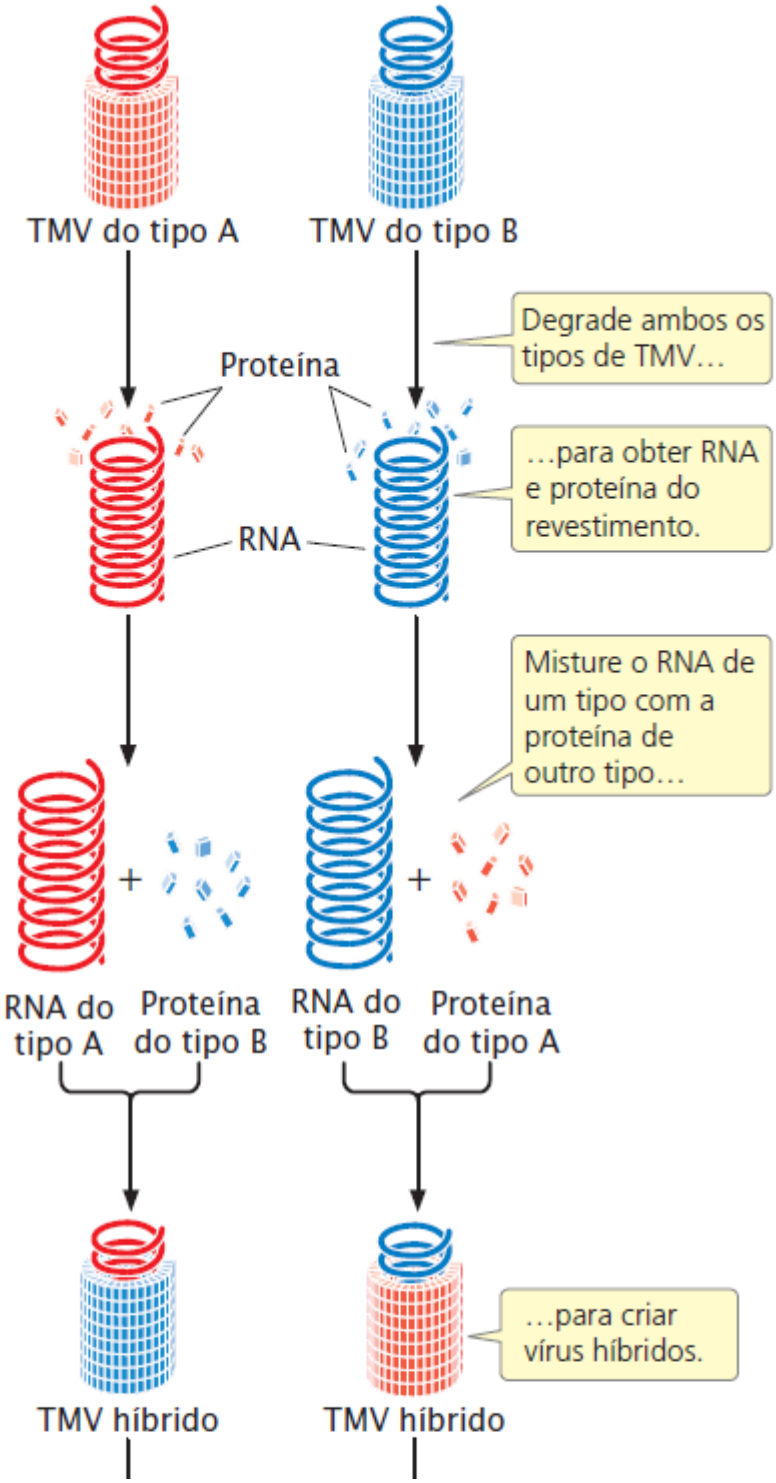
seu material genético. Essa modalidade foi demonstrada em 1956 por Heinz Fraenkel-Conrat e Bea Singer, que trabalharam com o vírus do mosaico do tabaco (TMV), um vírus que infecta e provoca doença nas plantas do tabaco (**Figura 10.8**). O TMV tem uma única molécula de RNA circundada por um cilindro helicoidal de moléculas de proteína. Fraenkel-Conrat descobriu que, após separar o RNA e a proteína do TMV, ele poderia remisturar o RNA e a proteína de diferentes cepas de TMV e obter partículas virais intactas, infectantes.

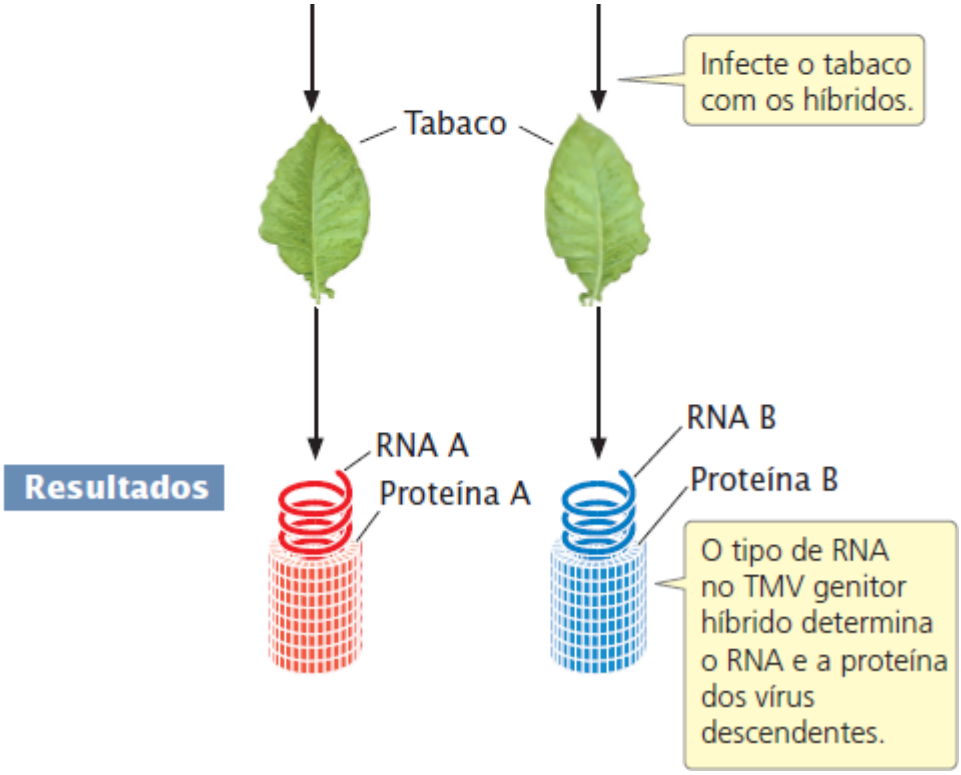
Com Singer, Fraenkel-Conrat criou vírus híbridos ao misturar o RNA e a proteína de diferentes cepas de TMV. Quando esses vírus híbridos infectavam as folhas de tabaco, novas partículas virais eram produzidas. Os novos descendentes virais eram idênticos à cepa a partir da qual foi isolado o RNA e não exibiam as características da cepa que doou a proteína. Esses resultados mostraram que o RNA carreia a informação genética no TMV.

Experimento

Pergunta: Que substância – RNA ou proteína – carreia o material genético no vírus do mosaico do tabaco (TMV)?

Métodos





Conclusão: RNA é o material genético do TMV.

Figura 10.8 O experimento de Fraenkel-Conrat e Singer demonstrou que o RNA no vírus do mosaico do tabaco carreia as informações genéticas.

Ainda em 1956, Alfred Gierer e Gerhard Schramm demonstraram que apenas o RNA isolado do TMV era suficiente para infectar as plantas do tabaco e direcionar a produção de novas partículas de TMV. Esse achado confirmou que o RNA carreia as instruções genéticas. Resolva o Problema 18

Conceitos

O RNA serve como material genético em alguns vírus.

10.3 O DNA consiste em duas cadeias de nucleotídios complementares e antiparalelas que formam uma dupla-hélice

O DNA, embora apresente uma estrutura simples, tem uma elegância e beleza insuperável por outras grandes moléculas. É útil analisar a estrutura do DNA nos três níveis de complexidade crescente, conhecidos como estruturas primária, secundária e terciária do DNA. A estrutura primária refere-se à estrutura de nucleotídios e como eles são unidos. A segunda estrutura refere-se à configuração tridimensional estável do DNA, a estrutura helicoidal descoberta por Watson e Crick. No Capítulo 11, vamos analisar as estruturas terciárias do DNA, que são arranjos de empacotamento complexo do DNA de fita dupla nos cromossomos.

Estrutura primária do DNA

A estrutura primária do DNA consiste em uma cadeia de nucleotídios unidos por meio de ligações fosfodiéster.

Nucleotídios. O DNA é uma molécula muito longa, chamada de macromolécula. Por exemplo, em cada cromossomo humano há uma única molécula de DNA, que, se alongada, teria centímetros de comprimento, milhares de vezes mais longa que a célula. Apesar do seu tamanho, o DNA tem uma estrutura muito simples: é um polímero, ou seja, uma cadeia feita com muitas unidades repetidas unidas. As unidades repetidas do DNA são os *nucleotídios*, cada um composto por três partes: (1) um açúcar, (2) um fosfato, e (3) uma base com nitrogênio.

Os açúcares dos ácidos nucleicos – chamados pentose – têm cinco átomos de carbono, numerados 1′, 2′, 3′ e assim por diante (**Figura 10.9**). Os açúcares do DNA e do RNA apresentam estrutura um pouco diferente. O açúcar do RNA, chamado **ribose**,

tem um grupo hidroxila (—OH) preso ao átomo do carbono 2′, enquanto o açúcar do DNA, ou **desoxirribose**, tem um átomo de hidrogênio (—H) nessa posição e, portanto, um átomo de oxigênio a menos. Essa diferença originou os nomes ácido ribonucleico (RNA) e ácido *desoxi*rribonucleico (DNA). Essa pequena diferença química é identificada pela maioria das enzimas celulares que interagem com DNA e RNA, gerando assim funções específicas para cada ácido nucleico. Além disso, o átomo de oxigênio adicional no nucleotídio do RNA o torna mais reativo e menos estável do ponto de vista químico, em comparação com o DNA. Por esse motivo, o DNA se encaixa melhor como repositório a longo prazo de informação genética.

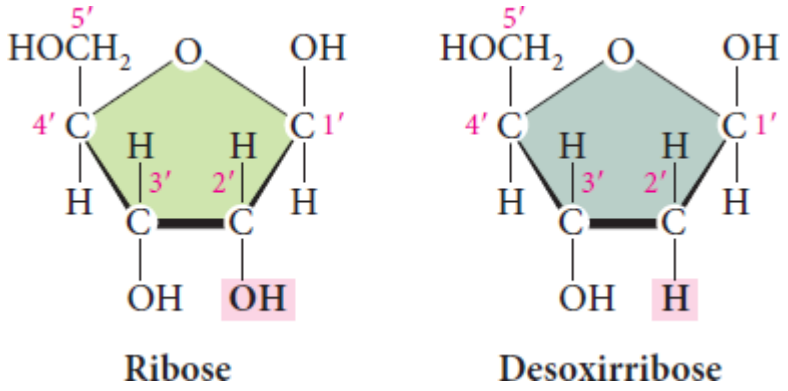


Figura 10.9 Um nucleotídio tem um açúcar ribose (RNA) ou um açúcar desoxirribose (DNA). Os átomos de carbono recebem números apostrofados

O segundo componente de um nucleotídio é sua **base nitrogenada**, que pode ser uma **purina** ou uma **pirimidina** (**Figura 10.10**). Cada purina tem um anel de seis membros preso a um anel de cinco membros, enquanto cada pirimidina tem apenas um anel de seis membros. O DNA e o RNA têm as duas purinas, **adenina** e **guanina** (A e G), com diferença nas posições de suas duplas ligações e nos grupos presos ao anel de seis membros. Três pirimidinas são comuns nos ácidos nucleicos: **citosina** (C), **timina** (T) e **uracila** (U). A citosina é encontrada no DNA e no RNA; entretanto, a timina está restrita ao DNA, e a uracila é encontrada apenas no RNA. As três pirimidinas apresentam diferenças nos grupos ou átomos presos aos átomos de carbono e no número de ligações duplas no anel. Em um nucleotídio, a base nitrogenada sempre forma uma ligação covalente com o átomo 1′ de carbono do açúcar (Figura 10.9). Um açúcar desoxirribose ou ribose e uma base juntos são chamados de **nucleosídio**.

O terceiro componente de um nucleotídio é o **grupo fosfato**, que consiste em um átomo de fósforo ligado a quatro átomos de oxigênio (**Figura 10.11**). Os grupos de fosfato são encontrados em todos os nucleotídios e têm carga elétrica negativa, tornando o DNA ácido. O grupo fosfato está sempre ligado ao átomo 5′ de carbono do açúcar (ver Figura 10.9) em um nucleotídio.

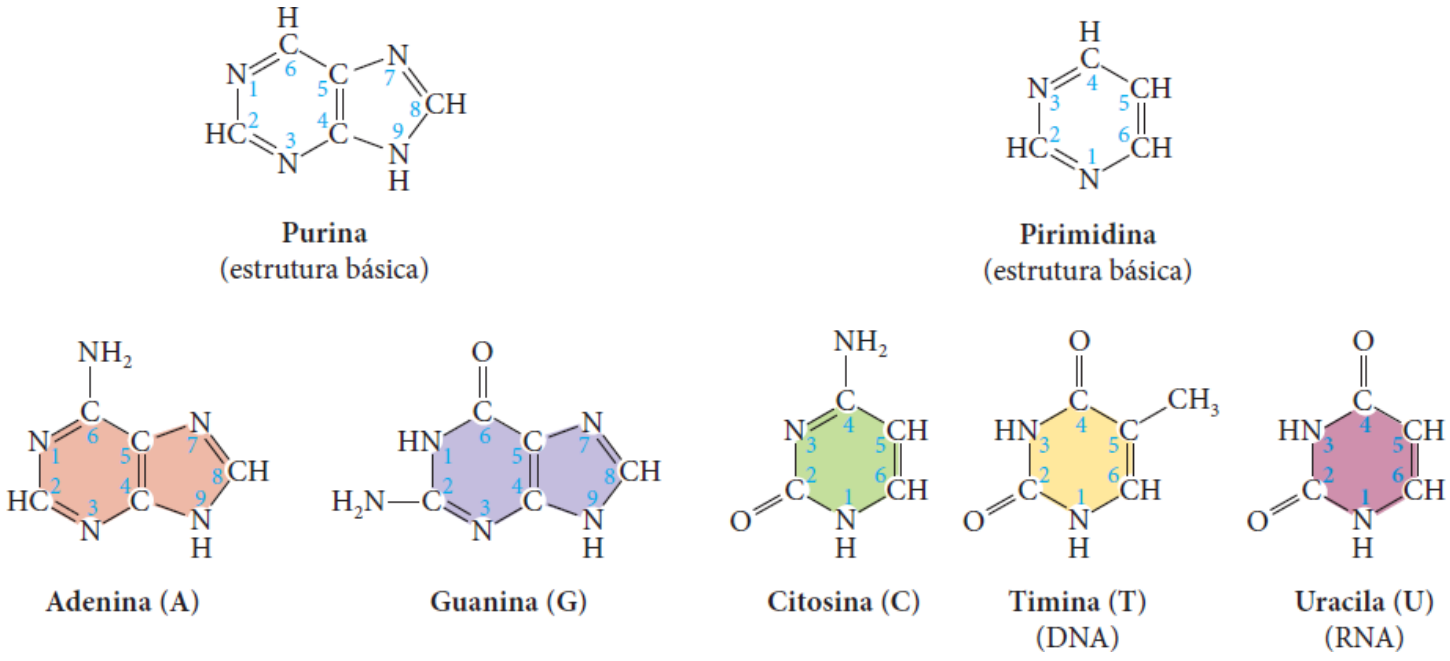


Figura 10.10 Um nucleotídio tem uma base purina ou pirimidina. Os átomos dos anéis nas bases recebem números não apostrofados.

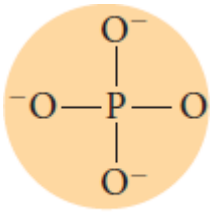


Figura 10.11 **O nucleotídio tem um grupo fosfato.**

Os nucleotídios do DNA são conhecidos como **desoxirribonucleotídios** ou 5′-monofosfatos de desoxirribonucleosídio. Como há quatro tipos de bases, há também quatro tipos diferentes de nucleotídios do DNA (**Figura 10.12**). Os nucleotídios RNA equivalentes são chamados **ribonucleotídios** ou 5′-monofosfatos de ribonucleosídio. Às vezes as moléculas de DNA têm bases raras adicionais, que são as formas modificadas das quatro bases comuns. Essas bases modificadas serão discutidas com mais detalhes quando examinarmos a função das moléculas de RNA, no Capítulo 14. Os nomes para as bases de DNA, os nucleotídios e nucleosídios, estão apresentados no **Quadro 10.2**. Resolva o Problema 26

Conceitos

A estrutura primária do DNA consiste em uma cadeia de nucleotídios. Cada nucleotídio é constituído por um açúcar de cinco carbonos, um fosfato e uma base. Há dois tipos de bases de DNA, purinas (adenina e guanina) e pirimidinas (timina e citosina).

Checagem dos conceitos 6

Qual é a diferença entre os açúcares do DNA e do RNA?

a. O RNA tem um açúcar de seis carbonos; o DNA tem um açúcar de cinco carbonos.
b. O açúcar do RNA tem um grupo hidroxila que não é encontrado no açúcar do DNA.
c. O RNA tem uracila, o DNA tem timina.
d. O açúcar do DNA tem um átomo de fósforo; o açúcar do RNA não tem.

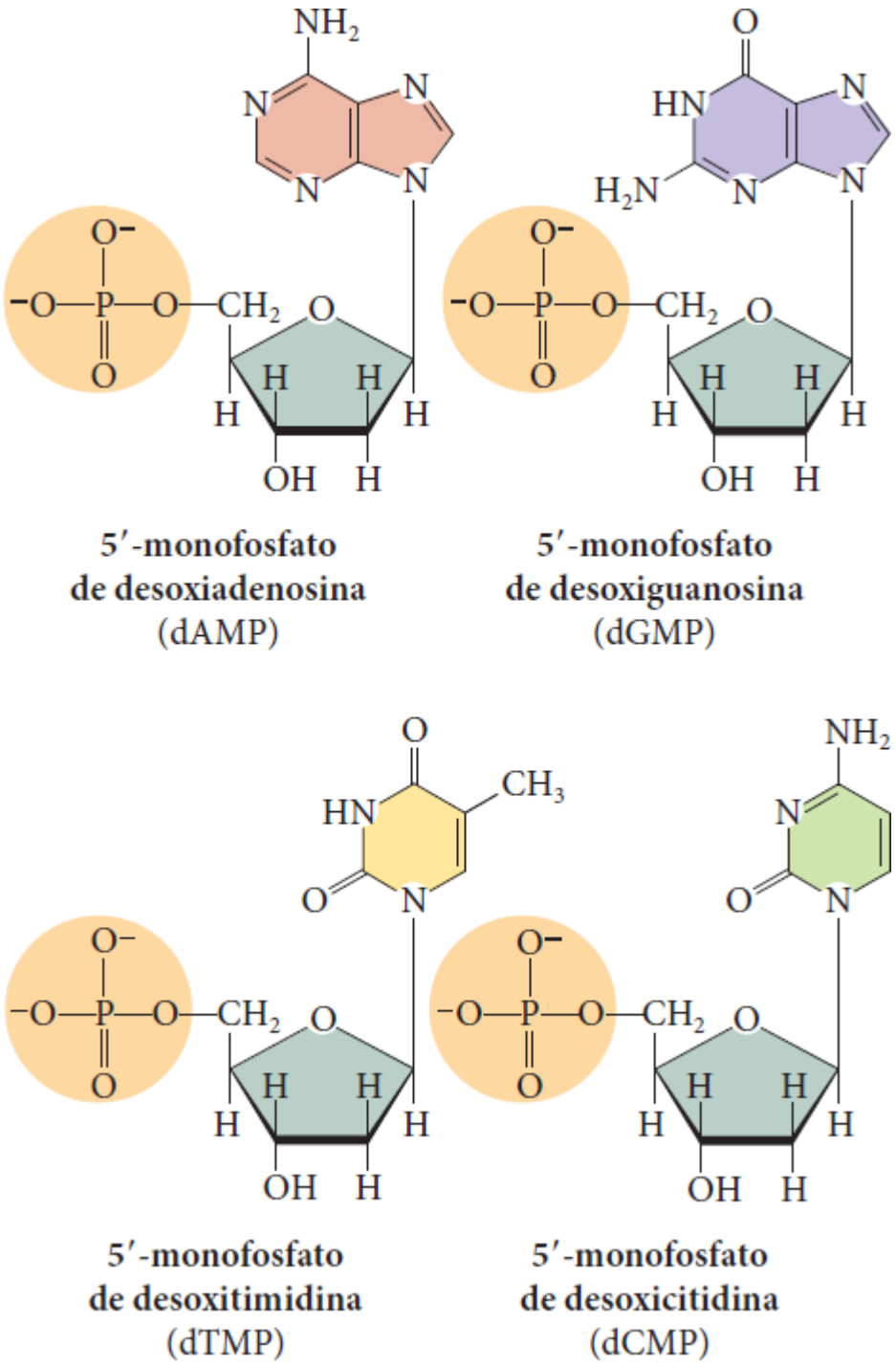


Figura 10.12 **Existem quatro tipos de nucleotídios do DNA.**

Cadeias de polinucleotídios. O DNA é constituído por muitos nucleotídios conectados por ligações covalentes, que se unem ao grupo 5′-fosfato de um nucleotídio com a 3′-hidroxila do próximo nucleotídio (**Figura 10.13**); observe que as estruturas apresentadas na Figura 10.13 estão restritas a duas dimensões enquanto a molécula é tridimensional, como demonstrado na Figura 10.14. Chamadas **ligações fosfodiéster**, elas são ligações covalentes fortes; vários nucleotídios ligados dessa forma constituem uma **cadeia de polinucleotídios**. O esqueleto da cadeia de polinucleotídios é composto de açúcares e fosfatos alternados; as bases se projetam a partir do longo eixo da cadeia. As cargas negativas dos grupos de fosfato são neutralizadas pela associação com as cargas positivas de proteínas, metais ou outras moléculas.

Uma importante característica da cadeia de polinucleotídios é sua direção, ou polaridade. Na extremidade da cadeia, um grupo fosfato livre (ou seja, que não está ligado em um dos lados) está preso ao átomo 5′-carbono do açúcar no nucleotídio. Essa extremidade da cadeia é chamada **extremidade 5′**. A outra extremidade da cadeia, a **extremidade 3′**, tem um grupo OH livre preso ao átomo 3′ de carbono do açúcar.

Os nucleotídios do RNA também estão conectados a ligações fosfodiéster para formar cadeias de polinucleotídios 5′ a 3′ semelhantes (ver Figura 10.13).

Conceitos

Os nucleotídios do DNA são unidos nas cadeias de polinucleotídios por ligações fosfodiéster que ligam o átomo 3′ de carbono de um nucleotídio ao grupo 5′-fosfato do nucleotídio seguinte. Cada cadeia de polinucleotídios tem polaridade, com uma extremidade 5′ e uma outra 3′.

Estruturas secundárias do DNA

A estrutura secundária do DNA refere-se à sua configuração tridimensional – sua estrutura helicoidal fundamental. Ela pode assumir várias configurações, dependendo de sua sequência de bases e das condições ambientais.

Dupla-hélice. Uma característica fundamental da estrutura secundária do DNA é que ele tem duas cadeias de polinucleotídios enroladas uma sobre a outra – é uma *dupla*-hélice. As ligações açúcar-fosfato estão no lado externo da hélice, e as bases estão empilhadas no interior da molécula (ver Figura 10.13). As duas cadeias de polinucleotídios seguem em direções opostas – são **antiparalelas**; isso significa que a extremidade 5′ de uma cadeia está oposta à extremidade 3′ da outra.

As cadeias são mantidas unidas por dois tipos de forças moleculares. As pontes de hidrogênio ligam as bases nas cadeias opostas (ver Figura 10.13). Essas ligações são relativamente fracas se comparadas com as ligações covalentes fosfodiéster que conectam o açúcar e os grupos de fosfato dos nucleotídios vizinhos na mesma cadeia. Como veremos adiante, várias funções importantes do DNA demandam a separação de suas duas cadeias de nucleotídios, e é fácil separá-las por causa da relativa facilidade em romper e restabelecer as pontes de hidrogênio.

A natureza da ponte de hidrogênio impõe uma limitação nos tipos de base que podem parear. Normalmente a adenina pareia apenas com a timina com duas pontes de hidrogênio, e a citosina pareia apenas com a guanina com três pontes de hidrogênio (ver Figura 10.13). Como se formam três pontes de hidrogênio entre C e G e apenas duas entre A e T, o pareamento C-G é mais forte que o A-T. A especificidade do pareamento das bases significa que, onde quer que exista uma A em uma cadeia, deve existir uma T na posição correspondente na outra cadeia, e onde quer que exista uma G em uma cadeia, deve existir uma C no outro. As duas fitas de polinucleotídios de uma molécula de DNA, portanto, não são idênticas, mas sim **fitas complementares do DNA**. A natureza complementar das duas cadeias de nucleotídios possibilita uma replicação eficiente e precisa do DNA, como será discutido no Capítulo 12.

Quadro 10.2 Nomes das bases, nucleotídios e nucleosídios do DNA.

	Adenina	Guanina	Timina	Citosina
Símbolo da base	A	G	T	C
Nucleotídio	5′-monofosfato de desoxiadenosina	5′-monofosfato de desoxiguanosina	5′-monofosfato de desoxitimidina	5′-monofosfato de desoxicitidina
Símbolo do nucleotídio	dAMP	dGMP	dTMP	dCMP
Nucleosídio	desoxiadenosina	desoxiguanosina	desoxitimidina	desoxicitidina
Símbolo do nucleosídio	dA	dG	dT	dC

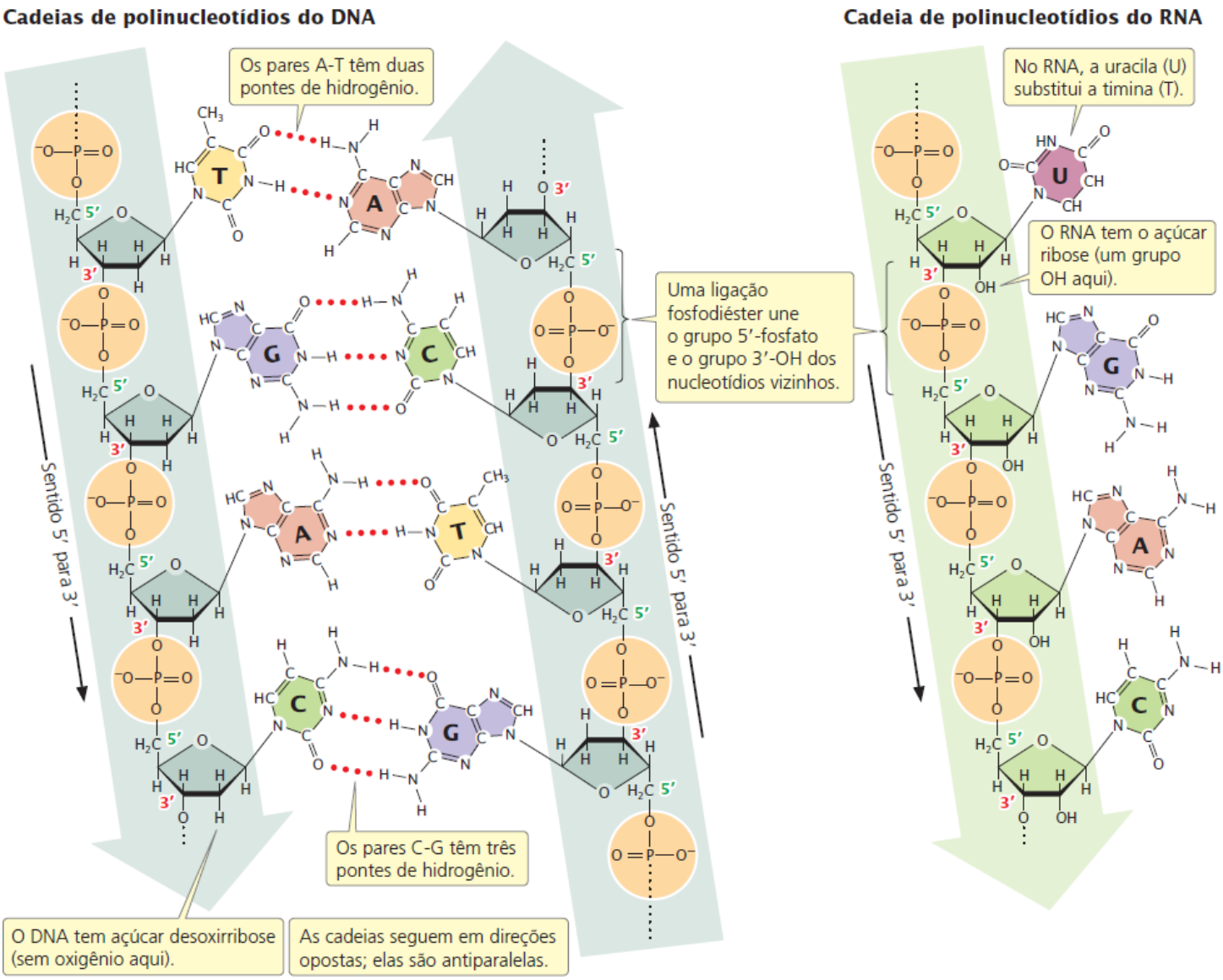


Figura 10.13 DNA e RNA são compostos por cadeias de polinucleotídios. O DNA é composto por duas cadeias de polinucleotídios, embora DNA de fita única seja encontrado em alguns vírus.

A segunda força que mantém duas cadeias de DNA unidas é a interação dos pares de bases empilhadas no interior da molécula. O empilhamento significa que as bases adjacentes estão alinhadas de modo que seus anéis estão paralelos e empilhados um em cima do outro. As interações de empilhamento estabilizam a molécula de DNA, mas não exigem que uma base particular acompanhe outra. Assim, a sequência de bases da molécula de DNA tem liberdade para variar, permitindo que o DNA carreie a informação genética. Resolva os Problemas 32 e 39

Conceitos

O DNA consiste em duas cadeias de polinucleotídios. Os grupos açúcar-fosfato de cada cadeia de polinucleotídios estão fora da molécula, e as bases estão no interior. A ponte de hidrogênio une as bases das duas cadeias: guanina pareia com citosina e adenina pareia com timina. As duas cadeias de polinucleotídios de uma molécula de DNA são complementares e antiparalelas.

Checagem dos conceitos 7

A natureza antiparalela do DNA refere-se:

a. A seus grupos de fosfato com carga elétrica.
b. Ao pareamento das bases em uma cadeia com as bases de outra cadeia.
c. À formação de pontes de hidrogênio entre as bases de fitas opostas.
d. Ao sentido oposto das duas cadeias de nucleotídios.

Diferentes estruturas secundárias. Como observamos, o DNA tem duas cadeias de polinucleotídios antiparalelas e

complementares (as exceções são as moléculas de DNA de fita única em alguns vírus). O formato tridimensional preciso da molécula pode variar, entretanto, dependendo das condições em que o DNA esteja e, em alguns casos, da própria sequência de bases.

A estrutura tridimensional do DNA descrita por Watson e Crick é chamada de estrutura **B-DNA** (**Figura 10.14**). Essa estrutura existe quando a molécula está circundada por muita água e não há uma sequência de bases incomum no DNA – condições que ocorrem nas células. A estrutura B-DNA é a configuração mais estável para uma sequência aleatória de nucleotídios sob condições fisiológicas, e outros indícios sugerem que é a estrutura predominante na célula.

O B-DNA é uma hélice girando para a direita, o que significa que sua espiral está no sentido horário. Existem aproximadamente 10 pares de base (pb) a cada 360° de rotação da hélice, então cada par de bases está torcido 36° em relação às bases adjacentes (ver Figura 10.14 B). Esses pares de base estão distantes 0,34 nanômetro (nm), então cada rotação completa da molécula mede 3,4 nm. O diâmetro da hélice é 2 nm, e as bases estão perpendiculares ao longo do eixo da molécula de DNA. Um modelo de preenchimento de espaço mostra que o B-DNA tem uma estrutura fina e alongada (ver Figura 10.14 A). A espiral das cadeias de nucleotídios cria sulcos maiores e menores na hélice, características importantes para a ligação de algumas proteínas que regulam a expressão da informação genética (ver *Capítulo 16*).

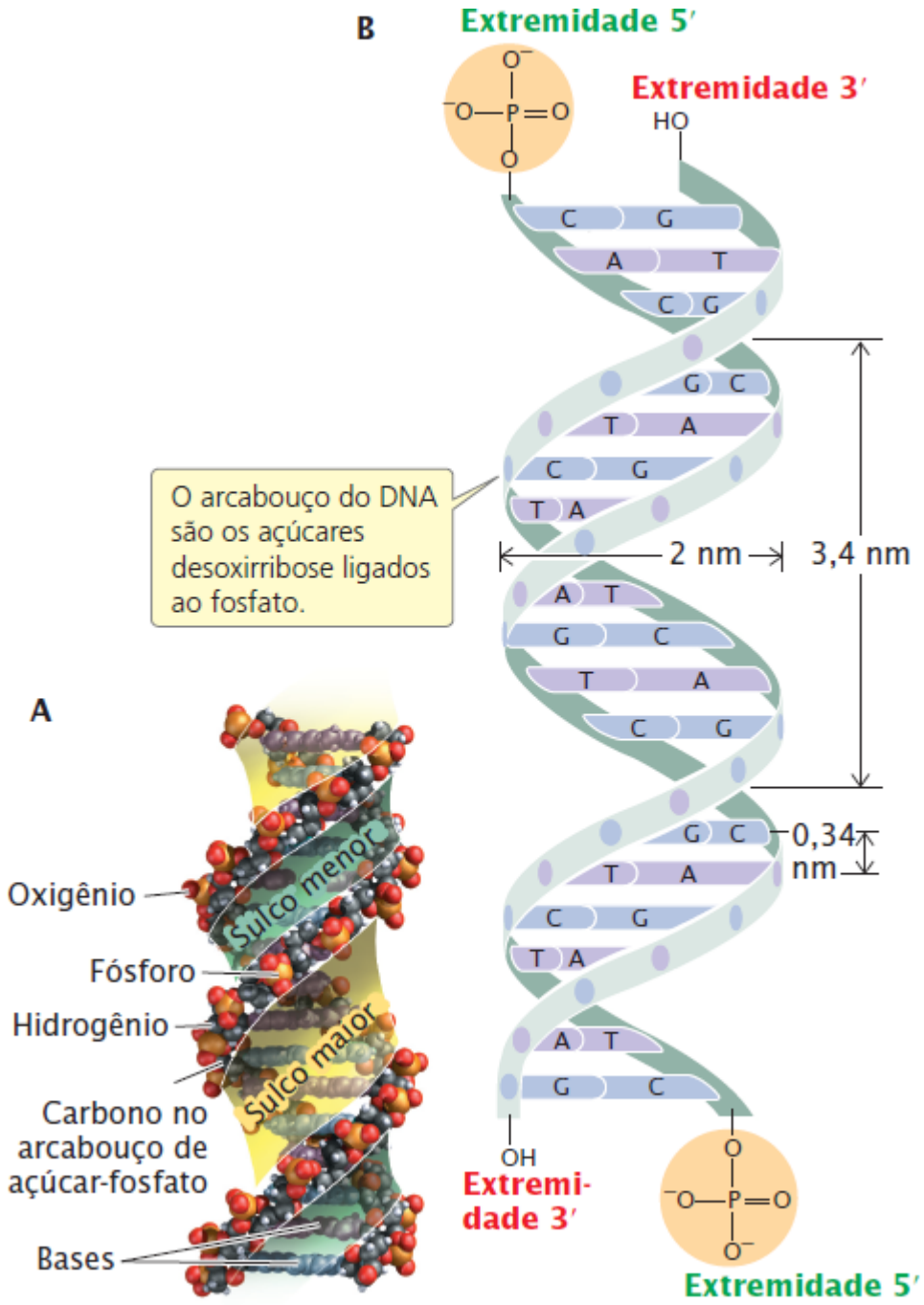


Figura 10.14 O B-DNA tem uma hélice girando para a direita com aproximadamente 10 bases por giro. A. Modelo de preenchimento espacial do B-DNA mostrando os sulcos maiores e menores. **B.** Representação em diagrama.

Outra estrutura secundária que o DNA pode assumir é a estrutura **A-DNA**, que existe no ambiente com pouca água. Assim

como o B-DNA, o A-DNA é uma hélice girando para a direita (**Figura 10.15 A**), mas é mais curta e mais larga que o B-DNA (**Figura 10.15 B**) e suas bases estão inclinadas para fora do eixo principal da molécula. O A-DNA foi detectado em alguns complexos de DNA-proteína e em esporos de algumas bactérias.

Uma estrutura secundária radicalmente diferente, chamada **Z-DNA** (**Figura 10.15 C**), forma uma hélice girando para a esquerda. Nessa forma, o esqueleto de açúcar-fosfato faz um zigue-zague para frente e para trás, dando origem ao seu nome. Uma estrutura Z-DNA pode surgir se a molécula tem uma sequência de bases particular, como os trechos de nucleotídios C e G alternados. Pesquisadores descobriram que anticorpos específicos para Z-DNA se ligam a regiões do DNA que foram transcritas em RNA, sugerindo que o Z-DNA participa na expressão do gene. Outras estruturas secundárias adicionais do DNA (C-DNA, D-DNA etc.) podem surgir sob condições laboratoriais específicas ou no DNA com sequências de base específicas.

Conceitos

O DNA pode assumir diferentes estruturas secundárias, dependendo das condições em que ele foi colocado e sua sequência de bases. Acredita-se que o B-DNA seja a configuração mais comum na célula.

Checagem dos conceitos 8

Qual é a diferença entre Z-DNA e B-DNA?

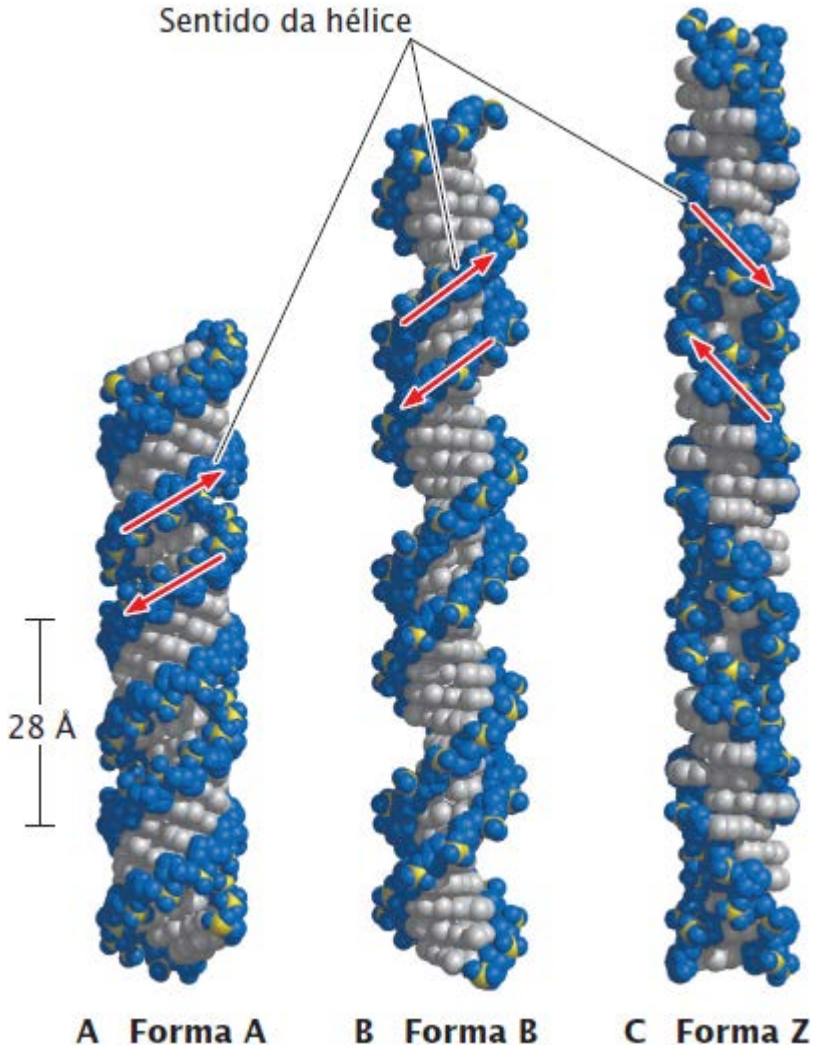


Figura 10.15 O DNA pode assumir várias estruturas secundárias diferentes. (De J. M. Berg, J. L. Tymoczko, and L. Stryer, Biochemistry, 6th ed. [New York: W. H. Freeman and Company, 2002], pp. 785 and 787.)

Conceitos conectantes

Implicações genéticas da estrutura do DNA

A imensa contribuição de Watson e Crick foi sua explicação sobre a estrutura química do genótipo, possibilitando que os geneticistas começassem a examinar os genes diretamente, em vez de olhar apenas

para as consequências fenotípicas da ação do gene. A determinação da estrutura do DNA levou ao nascimento da genética molecular – o estudo das naturezas química e molecular da informação genética.

A estrutura de Watson e Crick foi além de apenas criar o potencial para os estudos de genética molecular; ela foi uma fonte imediata de discernimento para os processos genéticos críticos. O início deste capítulo identificou quatro propriedades fundamentais do material genético. Primeiro, ele tem de ser capaz de carrear muita informação. O modelo de Watson e Crick sugeriu que as instruções genéticas estão codificadas na sequência das bases, a única parte variável da molécula.

Uma segunda propriedade necessária do material genético é sua capacidade de se replicar com segurança. As cadeias de polinucleotídios complementares do DNA tornam essa replicação possível. Watson e Crick propuseram que na replicação, duas cadeias de polinucleotídios se descompactam, rompendo as fracas pontes de hidrogênio entre as duas fitas e cada fita serve como um molde no qual uma nova fita é sintetizada. A especificidade do pareamento das bases significa que apenas uma possível sequência de bases, a sequência complementar, pode ser sintetizada a partir de cada molde. Novas moléculas de DNA de fita dupla recém-replicadas serão idênticas à molécula de DNA de fita dupla original (ver *Capítulo 12*, sobre replicação de DNA).

Uma terceira propriedade essencial do material genético é a capacidade de expressar suas instruções no fenótipo. O DNA expressa suas instruções genéticas ao transferir suas informações para uma molécula de RNA, em um processo chamado **transcrição** (ver *Capítulo 13*). O termo *transcrição* é apropriado porque, embora a informação seja transferida do DNA para o RNA, ela permanece na linguagem dos ácidos nucleicos. Em alguns casos, a molécula de RNA transfere a informação genética para uma proteína ao indicar sua sequência de aminoácidos. Esse processo é chamado **tradução** (ver *Capítulo 15*), porque a informação tem de ser *traduzida* a partir da linguagem dos nucleotídios para a linguagem dos aminoácidos. Uma quarta propriedade do DNA é sua capacidade de variar. Essa variação consiste em diferenças na sequência de bases encontrada em diferentes indivíduos.

Agora podemos identificar as três principais vias de fluxo de informação na célula (**Figura 10.16 A**): na **replicação**, a informação passa de uma molécula de DNA para outra molécula de DNA; na transcrição, a informação passa do DNA para o RNA; e, na tradução, a informação passa do RNA para a proteína. Esse conceito de fluxo de informação foi formalizado por Francis Crick em um conceito que ele chamou de dogma central da biologia molecular. O **dogma central** afirma que a informação genética passa do DNA para a proteína em uma via unidirecional de informação de mão única. Agora, observamos, entretanto, que o dogma central é uma grande simplificação. Além das três vias gerais de informação de replicação, transcrição e tradução, podem ocorrer outras transferências em alguns organismos ou em circunstâncias especiais. Os retrovírus (ver *Capítulo 9*) e alguns elementos de transposição (ver *Capítulo 18*) transferem informação do RNA para o DNA (na **transcrição reversa**) e alguns vírus de RNA transferem informação de RNA para RNA (em **replicação de RNA**, **Figura 10.16 B**).

10.4 Estruturas especiais podem se formar no DNA e no RNA

Sequências *dentro* de uma única cadeia de nucleotídios podem ser complementares entre si e podem parear ao formar pontes de hidrogênio, produzindo regiões de fita dupla (**Figura 10.17**). Esse pareamento interno de base transmite uma estrutura secundária para uma molécula de fita única. Um tipo comum de estrutura secundária encontrada nas cadeias únicas de nucleotídios é o **grampo** (*hairpin*), que se forma quando sequências de nucleotídios na mesma cadeia são complementos invertidos (ver Figura 10.17 A). Um grampo é uma região de bases pareadas (a haste) e às vezes inclui bases não pareadas intercaladas (a alça). Quando as sequências complementares são contínuas, o grampo tem uma haste sem alça (ver Figura 10.17 B). As moléculas de RNA podem ter vários grampos, possibilitando que elas se dobrem em estruturas complexas (ver Figura 10.17 C). As estruturas secundárias são importantes nas funções de muitas moléculas de RNA, como veremos nos Capítulos 14 e 15.

As sequências de DNA também podem formar estruturas de três fitas (triplex), chamadas **H-DNA**, quando parte do DNA se desenrola e uma cadeia única de polinucleotídios de uma parte da molécula pareia com o DNA de fita dupla a partir de outra parte da molécula (**Figura 10.18**). Isso é possível porque, sob algumas condições, uma base pode parear simultaneamente com outras duas bases. O H-DNA ocorre em longas sequências contendo apenas bases de purina ou apenas bases de pirimidina. Algumas estruturas triplex têm uma cadeia de purinas pareada com duas cadeias de pirimidinas; outras estruturas triplex têm uma cadeia de pirimidina pareada com duas cadeias de purinas. As sequências capazes de adotar uma configuração H-DNA são comuns nos genomas dos mamíferos, e uma evidência sugere que o H-DNA ocorre em condições normais. Uma pesquisa

recente demonstrou que o H-DNA se degrada mais facilmente do que o DNA de fita dupla, o que então leva a maiores taxas de mutação onde ocorrem estruturas de H-DNA. As estruturas quadruplex envolvem quatro fitas de DNA e também podem surgir sob algumas condições. **Resolva o Problema 39**

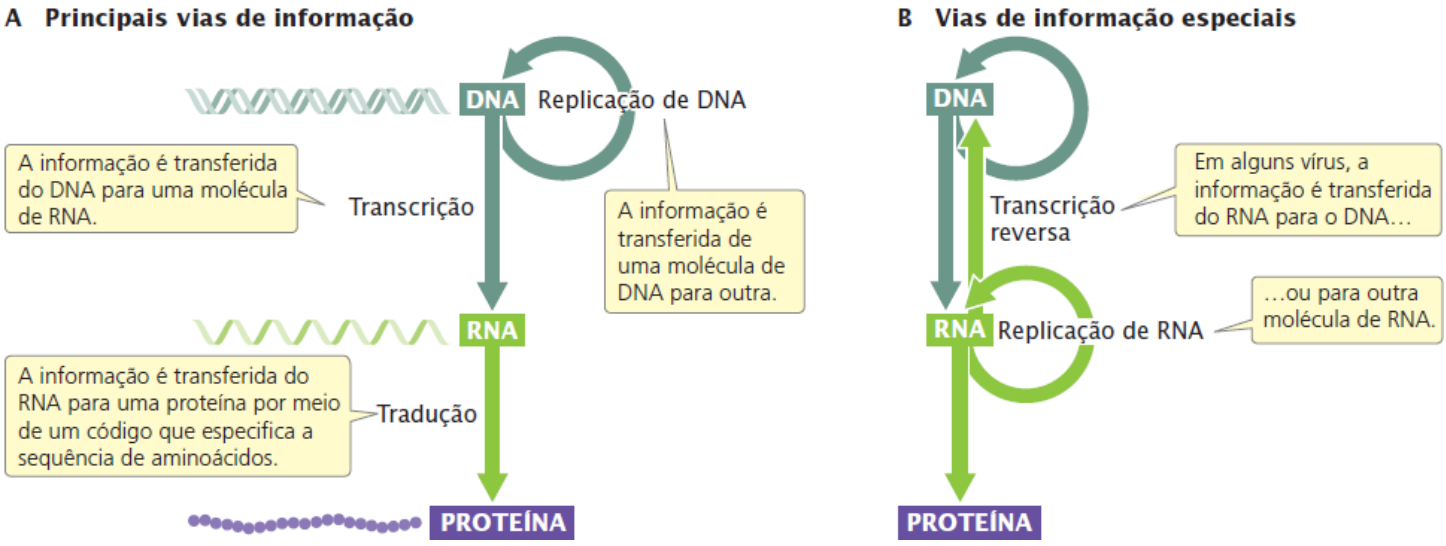


Figura 10.16 Vias de transferência de informação dentro da célula.

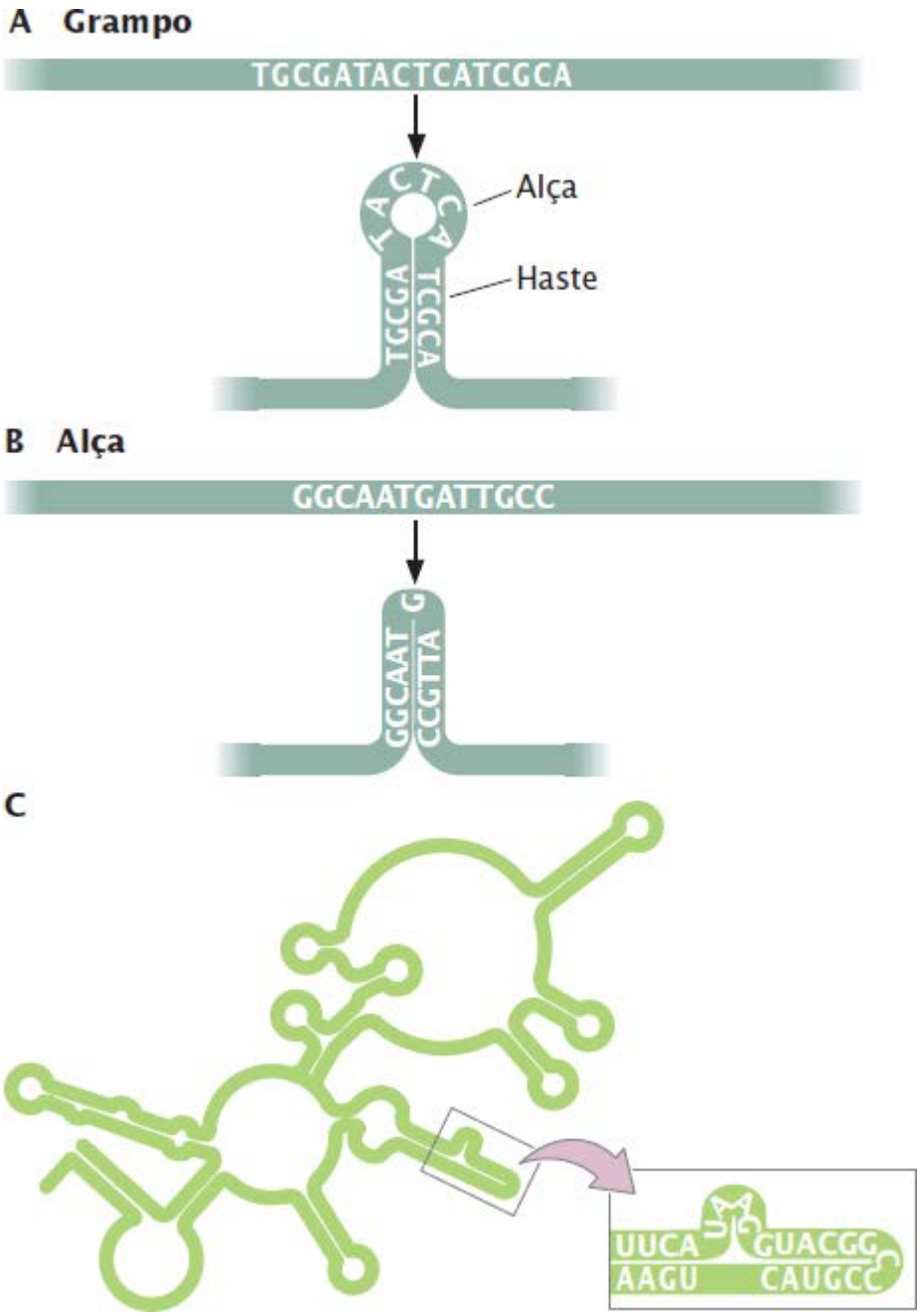


Figura 10.17 **DNA e RNA podem formar estruturas secundárias especiais. A.** Um grampo, consistindo em uma região de bases pareadas (que formam a haste) e uma região de bases não pareadas entre as sequências complementares (que formam uma alça na extremidade da haste). **B.** Uma haste sem alça. **C.** Estrutura secundária do componente RNA de RNase P de *E. coli.* As moléculas de RNA têm estruturas secundárias complexas.

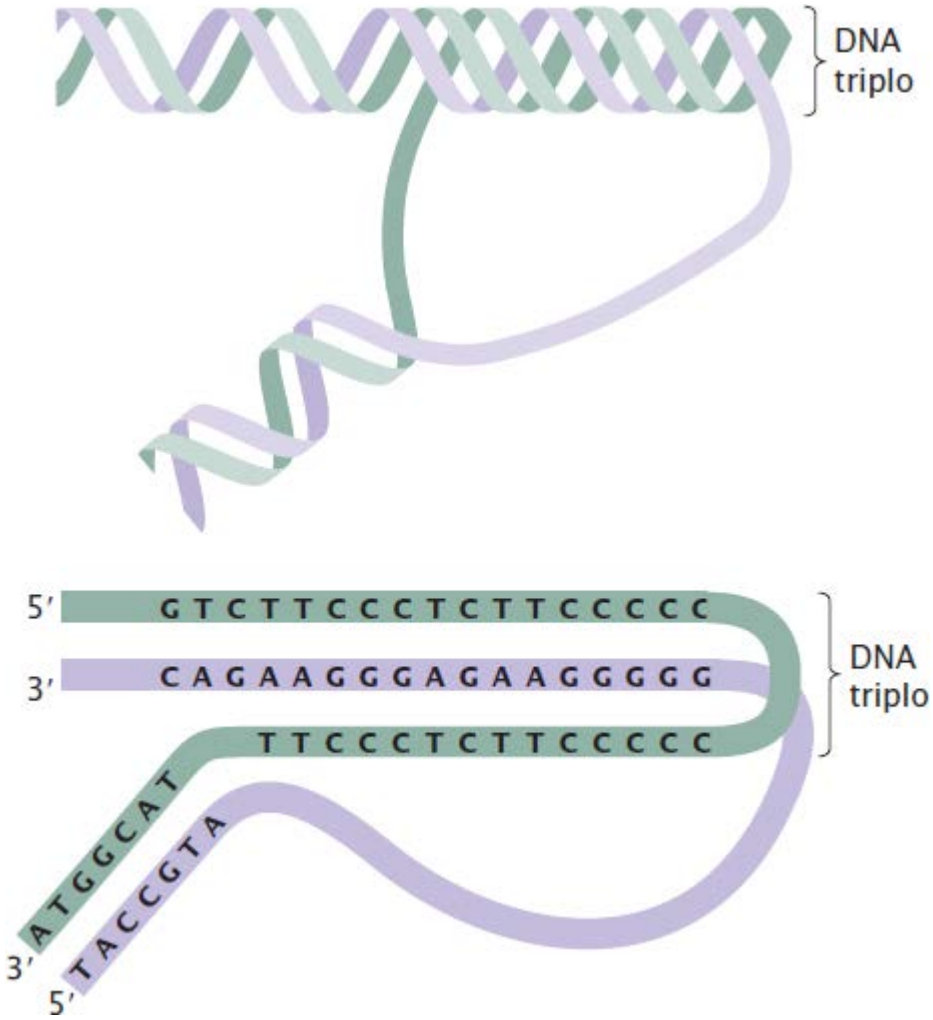


Figura 10.18 **O H-DNA surge quando três polinucleotídios pareiam.**

Conceitos

No DNA e no RNA, o pareamento de bases entre os nucleotídios na mesma fita produz estruturas secundárias especiais como grampos. As estruturas de DNA de fitas triplas podem surgir quando uma fita única de DNA pareia com DNA de fita dupla.

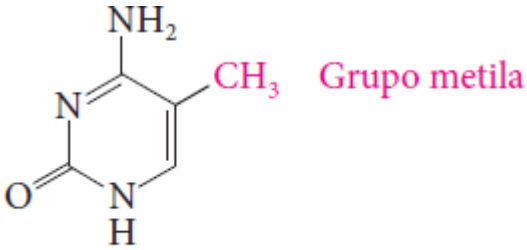
Checagem dos conceitos 9

Os grampos são formados no DNA como resultado de:

a. Sequências na mesma fita que estão invertidas e são complementares.
b. Sequências na fita oposta que são complementares.
c. Sequências na mesma fita que são idênticas.
d. Sequências na fita oposta que são idênticas.

A estrutura primária do DNA pode ser modificada de várias maneiras. Uma modificação é a **metilação do DNA**, um processo no qual os grupos metila (—CH_3) são adicionados (por enzimas específicas) para algumas posições nas bases dos nucleotídios. O DNA bacteriano é metilado para diferenciá-lo do DNA estranho, não metilado, que pode ser introduzido por vírus; as bactérias usam proteínas chamadas enzimas de restrição para cortar qualquer DNA viral não metilado (ver *Capítulo 19*). Nas células eucarióticas, a metilação está relacionada com a expressão do gene. As sequências metiladas apresentam baixos níveis de transcrição, enquanto as sequências sem metilação estão sendo transcritas de forma ativa (ver *Capítulo 17*). A metilação também pode afetar a estrutura tridimensional da molécula do DNA.

É comum encontrar adenina e citosina metiladas nas bactérias. No DNA eucariótico, as bases de citosina são metiladas às vezes para formar **5-metilcitosina** (**Figura 10.19**). A extensão da metilação da citosina varia entre os organismos eucarióticos; na maioria das células animais, cerca de 5% das bases de citosina estão metiladas, mas não existe citosina metilada na levedura e mais de 50% das bases de citosina estão metiladas em algumas plantas. Não está claro por que os organismos variam tanto em seu grau de metilação.



5-Metilcitosina

Figura 10.19 **No DNA eucariótico, as bases de citosina estão metiladas para formar 5-metilcitosina.**

Conceitos

Os grupos metila podem ser adicionados a algumas bases no DNA, dependendo da posição das bases na molécula. DNA procariótico e eucariótico podem estar metilados. Nos eucariotos, as bases de citosina estão mais metiladas para formar a 5-metilcitosina, e a metilação está frequentemente relacionada com a expressão do gene.

Resumo dos conceitos

- O material genético precisa conter informação complexa, ser replicado de forma exata, codificar o fenótipo e ter a capacidade de variar
- A evidência de que o DNA é a fonte de informação genética veio da descoberta de Avery, MacLeod e McCarty de que a transformação depende do DNA, e da demonstração de Hershey e Chase de que o DNA viral é transmitido para os fagos descendentes. Os resultados dos experimentos com TMV mostraram que o RNA carreia a informação genética em alguns vírus
- James Watson e Francis Crick, ao usar dados fornecidos por Rosalind Franklin e Maurice Wilkins, propuseram um modelo para a estrutura tridimensional do DNA em 1953
- Um nucleotídio de DNA é constituído por um açúcar desoxirribose, um grupo fosfato e uma base nitrogenada. O RNA é constituído por um açúcar ribose, um grupo fosfato e uma base nitrogenada
- Existem dois tipos de bases de um nucleotídio de DNA: purinas (adenina e guanina) e pirimidinas (citosina e timina). O RNA tem a pirimidina uracila em vez da timina
- Os nucleotídios são unidos por ligações fosfodiéster em uma cadeia de polinucleotídios. Cada cadeia de polinucleotídios tem um grupo fosfato livre na sua extremidade 5′ e um grupo hidroxila livre na extremidade 3′
- O DNA tem duas cadeias de nucleotídios que giram ao redor uma da outra para formar uma hélice dupla. Os açúcares e fosfatos ficam do lado de fora da hélice, e as bases estão presas no interior. As duas fitas são unidas por pontes de hidrogênio entre as bases de cada uma. As duas fitas são antiparalelas e complementares
- As moléculas de DNA podem assumir diferentes estruturas secundárias, dependendo das condições em que o DNA foi colocado e sua sequência de bases
- A estrutura do DNA tem muitas implicações genéticas importantes. A informação genética reside na sequência de bases do DNA, que por fim especifica a sequência de aminoácidos das proteínas. A complementariedade das bases nas duas fitas de DNA permite que a informação genética seja replicada
- O dogma central da biologia molecular propõe que a informação flui de modo unidirecional, do DNA para o RNA para a proteína. Atualmente, conhecemos exceções ao dogma central
- O pareamento entre as bases na mesma cadeia de nucleotídios pode levar à formação de grampos e a outras estruturas secundárias
- O DNA pode ser modificado com a adição de grupos metila às bases de nucleotídios.

Termos importantes

Adenina (A)
A-DNA
Antiparalela
Base nitrogenada

B-DNA
Cadeia de polinucleotídios
Citosina (C)
Desoxirribonucleotídio
Desoxirribose
Difração por raios X
Dogma central
Extremidade 3′
Extremidade 5′
Fitas complementares do DNA
Grampo
Grupo fosfato
Guanina (G)
H-DNA
Isótopos
Ligação fosfodiéster
Metilação do DNA
5-metilcitosina
Nucleosídio
Nucleotídio
Pirimidina
Princípio transformante
Purina
Regras de Chargaff
Replicação
Replicação de RNA
Ribonucleotídio
Ribose
Timina (T)
Tradução
Transcrição
Transcrição reversa
Uracila (U)
Z-DNA

Respostas da Checagem dos conceitos

1. Era impossível compreender como a informação genética era codificada sem conhecer a estrutura do DNA.
2. c.
3. c.
4. Não, porque o carbono é encontrado na proteína e no ácido nucleico.
5. d.
6. b.
7. d.
8. O Z-DNA tem uma hélice girando para a esquerda; B-DNA tem uma hélice girando para a direita. O esqueleto de açúcar-fosfato do Z-DNA faz um zigue-zague, enquanto o do B-DNA forma uma fita contínua e lisa.
9. a.

Problemas desenvolvidos

Problema 1

A porcentagem de citosina em uma molécula de DNA de fita dupla é 40%. Qual é a porcentagem de timina?

Estratégias para solução

De quais informações você precisa para solucionar o problema?
Da porcentagem de timina na molécula de DNA.

Quais informações são fornecidas para solucionar o problema?

- A molécula de DNA é fita dupla
- A porcentagem de citosina é 40%.

Para a solução deste problema, revise:
Estrutura primária do DNA e *Estruturas secundárias do DNA*, na Seção 10.3.

Etapas para a solução

Se C = 40%, então G também tem de ser 40%. A porcentagem total de C + G é, portanto 40% + 40% = 80%. Todas as bases remanescentes têm de ser A ou T; então, a porcentagem total de A +T = 100% − 80% = 20%. Como a porcentagem de A é igual à de T, a porcentagem de T é 20%/2 = 10%.

Lembrete: No DNA de fita dupla, A pareia com T, enquanto G pareia com C; então, a porcentagem de A é igual à de T; e a porcentagem de G é igual à de C.

Problema 2

Qual das relações a seguir é verdadeira para a porcentagem de bases no DNA de fita dupla?

a. C + T = A + G

b. $\frac{C}{A} = \frac{T}{G}$

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

Indicar se C + T = A + G e $\frac{C}{A} = \frac{T}{G}$ são verdadeiras.

Quais informações são fornecidas para solucionar o problema?

- O DNA é de fita dupla
- Razões de diferentes grupos de bases.

Para a solução deste problema, revise:
Estrutura primária do DNA e *Estruturas secundárias do DNA*, na Seção 10.3.

Etapas para a solução

Uma forma fácil de determinar se as relações são verdadeiras é indicar arbitrariamente porcentagens para as bases, lembrando que, no DNA de fita dupla, A = T e G = C. Por exemplo, se as porcentagens de A e T são 30% cada, então as porcentagens de G e C são 20% cada. Podemos substituir esses valores nas equações para ver se as relações são verdadeiras.

a. 20 + 30 = 30 + 20. Essa relação é verdadeira.
b. 20/30 ≠ 30/20. Essa relação não é verdadeira.

Questões de compreensão

Seção 10.1

1. Que quatro características gerais o material genético deve ter?

Seção 10.2

2. Descreva em poucas palavras a história do nosso conhecimento sobre a estrutura do DNA até o período de Watson e Crick. Quais foram as contribuições e os desenvolvimentos do princípio?
3. Que experimentos demonstraram que o DNA é o material genético?
4. O que é transformação? Como Avery e seus colegas demonstraram que o princípio transformante é o DNA?
5. Como Hershey e Chase mostraram que o DNA é transmitido para novos fagos na reprodução de fagos?
6. Por que a descoberta da estrutura do DNA foi tão importante?

Seção 10.3

7. Desenhe e identifique as três partes de um nucleotídio do DNA.
8. Qual é a diferença do nucleotídio do RNA para o nucleotídio do DNA?
9. Qual é a diferença da purina para a pirimidina? Quais purinas e pirimidinas são encontradas no DNA e no RNA?
10. Desenhe um pequeno segmento de fita única de polinucleotídios, incluindo pelo menos três nucleotídios. Indique a polaridade da fita ao identificar as extremidades 5′ e 3′.
11. Que bases são capazes de formar pontes de hidrogênio entre si?
12. Quais diferentes tipos de ligações químicas são encontrados no DNA e onde eles são encontrados?
13. Quais são algumas das implicações genéticas importantes da estrutura do DNA?
14. Quais são as três principais vias de fluxo de informação dentro da célula?

Seção 10.4

15. O que são grampos (*hairpin*) e como eles são formados?
16. O que é metilação do DNA?

Questões e problemas aplicados

Introdução

17. A introdução deste capítulo, sobre o sequenciamento do DNA de 4.000 anos, enfatiza a grande estabilidade do DNA. Quais aspectos da estrutura do DNA contribuem para a estabilidade da molécula? Por que o RNA é menos estável que o DNA?

Seção 10.2

*18. Combine os pesquisadores (a-j) com as descobertas listadas.
    a. Kossel
    b. Fraenkel-Conrat
    c. Watson e Crick
    d. Levene
    e. Miescher
    f. Hershey e Chase
    g. Avery, MacLeod e McCarty
    h. Griffith
    i. Franklin e Wilkins
    j. Chargaff

    ____ Tirou(aram) imagens com difração por raios X, usadas para construir a estrutura do DNA.
    ____ Determinou(aram) que o DNA tem bases nitrogenadas.
    ____ Identificou(aram) o DNA como material genético no bacteriófago.
    ____ Descobriu(ram) a regularidade nas razões de diferentes bases no DNA.
    ____ Determinou(aram) que o DNA é responsável pela transformação nas bactérias.
    ____ Solucionou(aram) a estrutura helicoidal do DNA ao construir modelos.
    ____ Descobriu(ram) que o DNA é composto por nucleotídios repetidos.
    ____ Determinou(aram) que o DNA é ácido e rico em fósforo.
    ____ Conduziu(ram) experimentos mostrando que o RNA pode servir como material genético em alguns vírus.
    ____ Demonstrou(aram) que o material de bactérias mortas pelo calor pode transformar geneticamente as bactérias vivas.

*19. Um estudante mistura bactérias *Streptococcus pneumoniae* do tipo IIS mortas pelo calor com bactérias do tipo IIR vivas e injeta a mistura em um camundongo. O camundongo desenvolve pneumonia e morre. O estudante recupera algumas bactérias do tipo IIS do animal morto. Se esse é o único experimento feito pelo estudante, ele teria demonstrado que ocorreu transformação? Quais outras explicações poderiam justificar as bactérias do tipo IIS no animal morto?

20. Imagine o que teria acontecido se Griffith tivesse misturado bactérias do tipo IIIS mortas pelo calor e algumas bactérias do tipo IIR mortas pelo calor e as injetasse em um camundongo. O animal teria contraído pneumonia e morrido? Explique.
21. Explique como as bactérias do tipo IIIS mortas pelo calor no experimento de Griffith alteraram geneticamente as bactérias do tipo IIR vivas. (Dica: veja a discussão sobre transformação no *Capítulo 9*.)
22. Que resultados você esperaria se o experimento de Hershey e Chase fosse feito no vírus do mosaico do tabaco?
23. Quais dos processos de transferência de informação ilustrados na Figura 10.16 são necessários para a reprodução do fago T2 ilustrada na Figura 10.4?

*24. Imagine que você é um estudante no laboratório de Alfred Hershey e Martha Chase no final de 1940. Você recebeu cinco tubos com a bactéria *E. coli* infectada pelo bacteriófago T2 que foi marcado com ^{32}P ou ^{35}S. Infelizmente, você se esqueceu de identificar os tubos e não tem certeza de quais estão marcados com ^{32}P e quais estão marcados com ^{35}S. Você coloca o conteúdo de cada tubo em um liquidificador e o liga por alguns segundos para romper os revestimentos de proteína. Então, centrifuga o conteúdo para separar os revestimentos de proteína e as células, e pesquisa sobre a existência de radioatividade, obtendo os resultados a seguir. Que tubos têm *E. coli* infectada com o fago marcado com ^{32}P? Explique sua resposta.

Número do tubo	Ocorrência de radioatividade em
1	Células
2	Revestimentos de proteína
3	Revestimentos de proteína
4	Células
5	Células

25. A Figura 10.8 ilustra o experimento de Fraenkel-Conrat e Singer sobre o material genético de TMV. Que resultados você esperaria nesse experimento se a proteína carreasse o material genético do TMV em vez do RNA?

Seção 10.3

*26. As moléculas de DNA de diferentes tamanhos são separadas com o uso de uma técnica chamada eletroforese (ver *Capítulo 19*). Com essa técnica, as moléculas de DNA são colocadas em um gel, uma corrente elétrica é aplicada ao gel e as moléculas de DNA migram no sentido do polo positivo (+) da corrente. Qual aspecto de sua estrutura faz uma molécula de DNA migrar no sentido do polo positivo?
*27. Cada par de nucleotídios de uma hélice dupla de DNA pesa cerca de 1×10^{-21} g. O corpo humano tem aproximadamente 0,5 g de DNA. Quantos pares de nucleotídios de DNA há no corpo humano? Se você considerar que todo DNA nas células humanas está na forma B-DNA, qual seria a distância necessária para esticar o DNA de uma extremidade a outra?
28. Uma cadeia de nucleotídios de molécula de DNA tem a sequência de bases ilustrada a seguir.

5′-ATTGCTACGG-3′

Indique a sequência de bases e marque as extremidades 5′ e 3′ da cadeia de nucleotídios complementar de DNA.

ANÁLISE DE DADOS

*29. Erwin Chargaff coletou dados sobre as proporções de bases de nucleotídios a partir do DNA de vários organismos e tecidos diferentes (E. Chargaff, in *The Nucleic Acids: Chemistry and Biology*, vol. 1, E. Chargaff and J. N. Davidson, Eds. New York: Academic Press, 1955). Os dados a seguir foram obtidos do DNA de vários organismos analisados por Chargaff.

	Porcentagem			
Organismo e tecido	**A**	**G**	**C**	**T**
Timo de carneiro	29,3	21,4	21,0	28,3
Fígado de porco	29,4	20,5	20,5	29,7
Timo humano	30,9	19,9	19,8	29,4
Medula óssea de rato	28,6	21,4	20,4	28,4
Eritrócitos de galinha	28,8	20,5	21,5	29,2

Levedura	31,7	18,3	17,4	32,6
E. coli	26,0	24,9	25,2	23,9
Espermatozoide humano	30,9	19,1	18,4	31,6
Espermatozoide de salmão	29,7	20,8	20,4	29,1
Espermatozoide de arenque	27,8	22,1	20,7	27,5

a. Para cada organismo, calcule a razão de (A + G)/(T + C) e a razão de (A + T)/(C + G).
b. Essas razões são constantes ou variam entre os organismos? Explique por quê.
c. A razão (A + G)/(T + C) é diferente nas amostras de espermatozoide? Você esperaria essa diferença? Por que sim ou por que não?

30. Boris Magasanik coletou dados sobre a quantidade de bases de DNA isolado de várias fontes, expressos relativos a um valor de 10 para adenina (B. Magasanik, in *The Nucleic Acids: Chemistry and Biology*, vol. 1, E. Chargaff and J. N. Davidson, Eds. New York: Academic Press, 1955).

	Porcentagem			
Organismo e tecido	**A**	**G**	**C**	**U**
Núcleos de fígado de rato	10	14,8	14,3	12,9
Núcleos de fígado de coelho	10	13,6	13,1	14,0
Cérebro de gato	10	14,7	12,0	9,5
Músculo de carpa	10	21,0	19,0	11,0
Levedura	10	12,0	8,0	9,8

a. Para cada organismo, calcule a razão (A + G)/(U + C).
b. Essas razões podem ser comparadas com a razão (A + G)/(T + C) encontrada no DNA (ver Problema 29)? Explique.

31. Quais relações ou razões a seguir seriam verdadeiras para uma molécula de DNA de fita dupla?

a. A + T = G + C
b. A + T = G + C
c. A + C = G + T
d. $\frac{A + T}{C + G} = 1{,}0$
e. $\frac{A + G}{C + T} = 1{,}0$
f. $\frac{A}{C} = \frac{G}{T}$
g. $\frac{A}{G} = \frac{C}{T}$
h. $\frac{A}{T} = \frac{G}{C}$

*32. Se uma molécula de DNA de fita dupla tem 15% de timina, quais são as porcentagens de todas as outras bases?

33. Suponha que cada uma das bases do DNA é capaz de parear com qualquer outra base. Que efeito esse fenômeno teria sobre a capacidade do DNA em servir como fonte de informação genética?

ANÁLISE DE DADOS

34. Heinz Shuster coletou os dados a seguir sobre a composição de bases do vírus *ribgrass* (H. Shuster, in *The Nucleic Acids: Chemistry and Biology*, vol. 3, E. Chargaff and J. N. Davidson, Eds. New York: Academic Press, 1955). Com base nessa informação, a informação hereditária do vírus *ribgrass* é o DNA ou o RNA? É provável que tenha fita única ou fita dupla?

Vírus do mosaico *ribgrass*.
(Jena Library of Biological Macromolecules.)

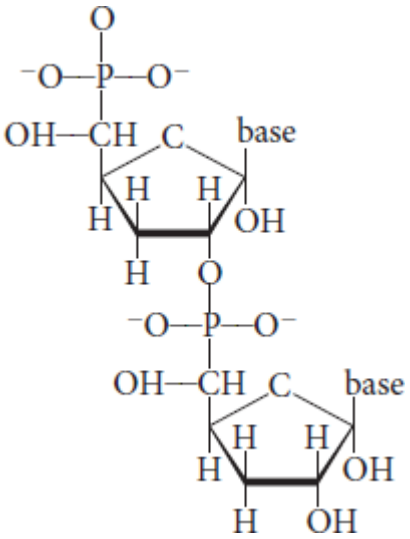
	Porcentagem				
	A	**G**	**C**	**T**	**U**
Vírus *ribgrass*	29,3	25,8	18,0	0,0	27,0

*35. As quantidades relativas de cada base de nucleotídio estão apresentadas no quadro a seguir para quatro diferentes vírus. Para cada vírus listado no quadro, indique se seu material genético é DNA ou RNA e se é fita única ou fita dupla. Apresente sua justificativa.

Vírus	**T**	**C**	**U**	**G**	**A**
I	0	12	9	12	9
II	23	16	0	16	23
III	34	42	0	18	39
IV	0	24	35	27	17

*36. Uma molécula B-DNA tem 1 milhão de pares de nucleotídios. Quantas voltas completas essa molécula tem?

*37. Como passatempo para a noite de sexta-feira, um professor de genética propôs que seus filhos fizessem um diagrama de uma cadeia de polinucleotídios de DNA. Como sua filha de 5 anos tinha aprendido sobre DNA na pré-escola, ela podia desenhar uma cadeia de polinucleotídios, mas cometeu alguns erros. O diagrama da filha (representado aqui) tinha pelo menos 10 erros.



a. Liste todos os erros na estrutura dessa cadeia de polinucleotídios de DNA.
b. Desenhe a estrutura da cadeia de polinucleotídios correta.

38. O Capítulo 1 analisou a teoria da herança de características adquiridas e observou que essa teoria não é mais aceita. O dogma central é consistente com a teoria da herança das características adquiridas? Por que sim ou por que não?

Seção 10.4

*39. Escreva uma sequência de bases em uma molécula de RNA que produzirá uma estrutura de grampo (*hairpin*).

40. Escreva uma sequência de nucleotídios em uma fita de DNA que formará uma estrutura de grampo (*hairpin*).

Questões desafiadoras

Seção 10.1

*41. Suponha que uma sonda automática, não tripulada, é enviada para o espaço profundo para pesquisa de vida extraterrestre. Após vagar por muitos anos-luz nos locais mais distantes do universo, essa sonda chega a um planeta distante e detecta vida. A composição química da vida nesse planeta é completamente diferente da composição da vida na Terra, e seu material genético não é composto por ácidos nucleicos. Que predições você pode tirar sobre as propriedades químicas do material genético desse planeta?

Seção 10.2

42. Como ^{32}P e ^{35}S poderiam ser usados para demonstrar que o princípio transformante é o DNA? Descreva em poucas palavras um experimento que mostraria que o DNA é o princípio transformante, em vez da proteína.

Seção 10.3

43. Pesquisadores propuseram que o início da vida na Terra usou RNA como fonte de informação genética e que posteriormente o DNA o substituiu. Quais aspectos da estrutura do DNA podem torná-lo mais adequado do que o RNA para ser o material genético?

44. Cientistas supostamente isolaram pequenos fragmentos de DNA de ossos de dinossauro fossilizado centenas de milhões de anos atrás. A técnica usada para isolar esse DNA é a reação em cadeia da polimerase, que é capaz de amplificar quantidades muito pequenas de DNA em milhares de vezes (ver *Capítulo 19*). Críticos afirmam que o DNA isolado de ossos de dinossauro não é oriundo apenas do dinossauro, mas foi contaminado por DNA de organismos atuais como bactérias, bolor ou seres humanos. Que precauções, análises e experimentos de controle poderiam ser feitos para garantir que o DNA recuperado dos fósseis é realmente antigo?

11

Estrutura do Cromossomo e DNA das Organelas

Telômeros e adversidade na infância

Cada uma de nossas células tem 46 cromossomos, estruturas primorosamente complexas de DNA e proteínas que carreiam as instruções da codificação de todos os nossos traços. Esses cromossomos são transmitidos por nossos pais e constituem a base da hereditariedade, a transmissão de traços de uma geração para a próxima. No entanto, os cromossomos não carreiam apenas um registro do nosso legado genético. Eles também são um registro, nas extensões de seus telômeros, dos estresses com os quais lidamos.

Telômeros são estruturas de proteção especiais encontradas na extremidade de cada um dos nossos cromossomos. Como pequenas pontas de plástico que impedem que os cadarços se descosturem, os telômeros evitam que as extremidades dos cromossomos sejam degradadas. Apesar dessa proteção, os cromossomos da maioria das células encurtam de forma progressiva em cada divisão celular. Em função de um "escape" da replicação do DNA, a maioria das células é incapaz de copiar todas as extremidades de cada cromossomo linear (ver *Capítulo 12* para uma discussão completa do problema da replicação da extremidade). Então, a cada rodada de replicação, um cromossomo se torna menor, até ficar tão reduzido que a célula para de se dividir, torna-se inativa e, consequentemente, morre. Para a maioria das células, esse encurtamento dos telômeros limita o número de divisões possíveis. Há exceções nas células germinativas que produzem as futuras gerações, algumas células-tronco e, infelizmente, em muitas células de câncer que escapam das restrições normais da divisão celular.

Uma vez que os telômeros se tornam menores a cada divisão celular, grande parte da pesquisa se concentra em determinar se o comprimento do telômero é indicativo de envelhecimento biológico. Embora a relação entre o comprimento do telômero e o envelhecimento seja complexa e ainda não esteja totalmente compreendida, indícios consideráveis sugerem que os telômeros encurtam com a idade, e que os processos que levam ao encurtamento prematuro do telômero estão associados às características do envelhecimento. Em 2011, os geneticistas observaram que as dificuldades encontradas no início da vida podem ter um papel no encurtamento dos nossos telômeros.

Para pesquisar os efeitos da experiência do início da vida sobre o comprimento do telômero, os geneticistas estudaram 100 crianças que viviam em orfanatos administrados pelo estado na Romênia. Em uma idade precoce, algumas dessas crianças foram adotadas e outras permaneceram nos orfanatos. Estudos anteriores demonstraram que as crianças nesses orfanatos receberam menos atenção e cuidado individual que as crianças que cresceram com os pais biológicos ou adotivos, e foi considerado que o cuidado institucional é mais estressante do que o cuidado dos pais adotivos.

Quando as crianças tinham de 6 a 10 anos, os pesquisadores coletaram amostras de seu DNA e mediram o comprimento de seus telômeros. Os resultados foram impressionantes: as crianças que permaneceram nos orfanatos apresentavam telômeros muito menores do que as que foram adotadas. Os pesquisadores concluíram que o comprimento do telômero é afetado pela adversidade na infância: as crianças criadas em ambientes estressantes apresentam maior probabilidade de terem telômeros menores do que as que cresceram em ambientes menos estressantes. Vários outros estudos encontraram uma associação semelhante entre o comprimento do telômero nos adultos e estresses no início da infância, como maus-tratos e doenças crônicas. Não sabemos como o estresse afeta os telômeros e leva ao seu encurtamento, mas a pesquisa documenta que os cromossomos são mais do que apenas um repositório de nossa informação genética – sua estrutura também é afetada pelo ambiente.

Este capítulo se dedica a examinar a estrutura molecular dos cromossomos e o DNA encontrado nas organelas citoplasmáticas. A primeira parte do capítulo se concentra no problema de armazenamento: como comprimir as imensas quantidades de DNA nos limites da fronteira de uma célula. Mesmo nos organismos com as menores quantidades de DNA, o comprimento do material genético excede muito o da célula. Assim, o DNA celular deve estar muito dobrado e condensado; essa condensação, porém, cria um problema: ela deixa o DNA inacessível, incapaz de ser copiado ou lido. O DNA funcional deve ser capaz de se abrir e se expandir parcialmente de modo que os genes individuais possam ser replicados e transcritos. A

natureza flexível e dinâmica da condensação do DNA é o principal assunto deste capítulo. Primeiro vamos analisar o superenrolamento, uma estrutura terciária importante do DNA encontrada nas células procarióticas e eucarióticas. Após um rápido exame no cromossomo bacteriano, examinaremos a estrutura dos cromossomos eucarióticos. Teremos atenção especial às partes funcionais de um cromossomo – especificamente, os centrômeros e telômeros. Também vamos analisar os tipos de sequências de DNA existentes em muitos cromossomos eucarióticos.

A segunda parte deste capítulo se volta para a organização das sequências de DNA encontrada nas mitocôndrias e nos cloroplastos. A herança uniparental exibida pelos genes encontrados nessas organelas foi discutida no Capítulo 5; aqui, examinaremos os aspectos moleculares do DNA dessas organelas. Analisaremos rapidamente as estruturas das mitocôndrias e dos cloroplastos, a herança dos traços codificados por seus genes e a origem evolutiva dessas organelas. Então, examinaremos as características gerais do **DNA mitocondrial** (mtDNA), seguido por uma discussão da organização e função dos diferentes tipos de genomas mitocondriais. Finalmente, vamos examinar o **DNA de cloroplasto** (cpDNA) e suas características, organização e função.

11.1 Muito DNA é condensado em uma célula

A condensação de grandes quantidades de informação genética no pequeno espaço de uma célula é chamada de problema fundamental de armazenamento. Considere o cromossomo da bactéria *E. coli*, uma única molécula de DNA com aproximadamente 4,6 milhões de pares de bases. Esticado em uma linha reta, esse DNA seria 1.000 vezes mais longo que a célula onde está localizado (**Figura 11.1**). As células humanas têm mais de 6 bilhões de pares de bases de DNA, que mediriam mais de 2 metros, esticados de uma ponta a outra. Até o DNA do menor cromossomo humano esticado seria 14.000 vezes o comprimento do núcleo. Está claro que as moléculas de DNA precisam ser bem comprimidas para se ajustarem em tais espaços pequenos.

A estrutura do DNA pode ser considerada em três níveis hierárquicos: a estrutura primária do DNA é sua cadeia de nucleotídios; a estrutura secundária é a hélice de fita dobrada; e a estrutura terciária refere-se à dobra de ordem superior que possibilita que o DNA seja acondicionado no espaço restrito de uma célula.

Conceitos

O DNA cromossômico existe na forma de moléculas muito longas que estão condensadas para se encaixar nos pequenos limites de uma célula.

Superenrolamento

Um tipo de estrutura terciária do DNA é o **superenrolamento**, que ocorre quando a hélice do DNA está sujeita à tensão ao ser supercontorcida ou subcontorcida. O menor estado de energia para o B-DNA (ver *Capítulo 10*) é quando ele tem aproximadamente 10 pb por volta de sua hélice. Nesse **estado relaxado do DNA**, uma distância de 100 pb de DNA teria cerca de 10 voltas completas. Se a energia é usada para adicionar ou remover qualquer volta, coloca-se tensão na molécula, fazendo com que a hélice superenrole ou gire sobre si mesma. Moléculas com rotação adicional exibem **superenrolamento positivo**. Moléculas com falta de rotação exibem **superenrolamento negativo**. O superenrolamento é uma solução parcial para o problema de condensação do DNA da célula porque o DNA superenrolado ocupa menos espaço que o DNA relaxado.

O superenrolamento ocorre quando a tensão da rotação não pode ser compensada ao girar as extremidades da dupla-hélice, que é o caso se o DNA for circular – ou seja, não existem extremidades livres. Se as cadeias *podem* girar livremente, suas extremidades simplesmente giram à medida que rotações extras são adicionadas ou removidas, e a molécula reverte espontaneamente para o estado relaxado. Tanto o DNA bacteriano quanto o eucariótico, em geral, dobram-se em alças estabilizados por proteínas (que evitam a rotação livre das extremidades, **Figura 11.2**), e o superenrolamento ocorre dentro das alças.

O superenrolamento depende das **topoisomerases**, enzimas que adicionam ou removem rotações da hélice do DNA ao romper temporariamente as fitas de nucleotídios, girando as extremidades ao redor de si mesmas e então restituindo as extremidades rompidas. Assim, as topoisomerases podem induzir ou aliviar o superenrolamento, embora nem todas as enzimas topoisomerases exerçam ambas as funções.

A maior parte do DNA encontrado nas células apresenta superenrolamento negativo, que tem duas vantagens sobre o DNA não superenrolado. Primeiro, o superenrolamento negativo facilita a separação das duas fitas de DNA durante a replicação e transcrição. O DNA com superenrolamento negativo tem falta de rotação, então a separação das duas fitas durante a replicação e a transcrição é mais rápida e requer menos energia. Segundo, o DNA superenrolado pode ser acondicionado em um espaço menor que o DNA relaxado.

Conceitos

A rotação adicional ou falta de rotação de uma dupla-hélice de DNA coloca tensão na molécula, fazendo com que ela superenrole. O superenrolamento é controlado por enzimas topoisomerases. A maior parte do DNA celular está negativamente superenrolada, o que facilita a separação das fitas de nucleotídios durante a replicação e a transcrição e possibilita o acondicionamento do DNA em espaços pequenos.

Checagem dos conceitos 1

Uma molécula de DNA de 300 pb tem 20 rotações completas. Essa molécula de DNA está:

a. Positivamente superenrolada.
b. Negativamente superenrolada.
c. Relaxada.

Figura 11.1 O DNA da *E. coli* é cerca de 1.000 vezes maior que a própria célula.

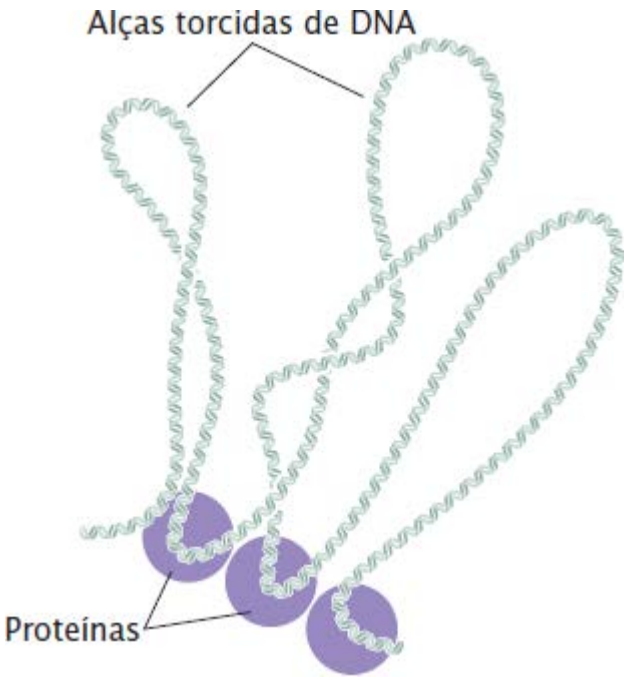


Figura 11.2 O DNA bacteriano está muito dobrado em uma série de alças torcidas.

Cromossomo bacteriano

A maioria dos genomas bacterianos é composta por uma única molécula de DNA circular, embora tenham sido encontradas moléculas de DNA lineares em algumas espécies. Nos cromossomos circulares bacterianos, o DNA não existe em um círculo aberto, relaxado; os 3 a 4 milhões de pares de base de DNA em um típico genoma bacteriano seriam muito grandes para se encaixar em uma célula bacteriana (ver Figura 11.1). O DNA bacteriano não está preso às histonas como o DNA eucariótico (discutido posteriormente neste capítulo), mas é complexado por várias proteínas que ajudam em sua compactação.

Quando uma célula bacteriana é visualizada com microscópio eletrônico, seu DNA aparece como um aglomerado distinto, o **nucleoide**, restrito a uma região definida do citoplasma. Se uma célula bacteriana é suavemente aberta, seu DNA vaza em uma série de alças torcidas. É mais provável que as extremidades das alças sejam mantidas no lugar pelas proteínas (ver Figura 11.2). Muitas bactérias exibem DNA adicional na forma de pequenas moléculas circulares chamadas de plasmídios, que se replicam de forma independente do cromossomo (ver *Capítulo 9*).

Conceitos

Um cromossomo bacteriano típico é composto por uma grande molécula circular de DNA que é uma série de alças torcidas. O DNA bacteriano aparece como um aglomerado distinto, o nucleoide, dentro da célula bacteriana.

Checagem dos conceitos 2

Qual é a diferença entre o DNA bacteriano e o DNA eucariótico?

Cromossomos eucarióticos

Os cromossomos eucarióticos contêm muito DNA. Como o cromossomo bacteriano, cada cromossomo eucariótico tem uma única molécula de DNA extremamente longa. Para que esse DNA se encaixe no núcleo, é necessário um enorme esforço de condensação e dobra, cuja intensidade varia durante o ciclo celular. Os cromossomos estão em um estado alongado, relativamente não condensado durante a interfase do ciclo celular (ver *Capítulo 2*), mas o termo *relativamente* é importante aqui. Embora o DNA dos cromossomos na interfase esteja menos condensado do que o DNA dos cromossomos mitóticos, ele ainda está muito condensado; está apenas *menos* condensado. No curso do ciclo celular, o nível de condensação do DNA muda: os cromossomos progridem de um estado altamente condensado para um estado de extrema condensação, necessário para o movimento dos cromossomos na mitose e meiose. A condensação do DNA também muda localmente na replicação e transcrição, quando as duas fitas de nucleotídios devem abrir e as sequências de bases em questão estão expostas. Assim, a condensação do DNA eucariótico (sua estrutura cromossômica terciária) não é estática – ela muda regularmente em resposta aos processos celulares.

Cromatina. O DNA eucariótico na célula está intimamente associado às proteínas. Essa combinação de DNA e proteínas é chamada de cromatina. Os dois tipos básicos de cromatina são a **eucromatina**, que sofre o processo normal de condensação e descondensação no ciclo celular, e a **heterocromatina**, que permanece em um estado altamente condensado durante todo o ciclo celular, mesmo durante a interfase. A eucromatina constitui a maior parte do material cromossômico e é onde grande parte da transcrição ocorre. Todos os cromossomos têm heterocromatina permanente (chamada de heterocromatina constitutiva) nos centrômeros e telômeros; o cromossomo Y também é composto, em grande parte, por heterocromatina constitutiva. A heterocromatina ainda pode ocorrer durante alguns estágios de desenvolvimento, e é então chamada de heterocromatina facultativa. Por exemplo, a heterocromatina facultativa ocorre ao longo de um cromossomo X inteiro nas fêmeas dos mamíferos quando esse X se torna inativo (ver *Capítulo 4*). Além de permanecer condensada durante todo o ciclo celular, a heterocromatina é caracterizada pela falta generalizada de transcrição, ausência de *crossing over* e replicação tardia na fase S. As diferenças entre a eucromatina e a heterocromatina estão resumidas no **Quadro 11.1**.

As proteínas mais abundantes na cromatina são as *histonas*, pequenas proteínas com carga elétrica positiva, com cinco tipos principais: H1, H2A, H2B, H3 e H4. Todas as histonas apresentam elevada porcentagem de arginina e lisina, aminoácidos com carga positiva que conferem às histonas uma carga elétrica efetiva positiva. As cargas positivas atraem as cargas elétricas negativas dos fosfatos do DNA; essa atração mantém o DNA em contato com as histonas. Uma porção heterogênea das **proteínas cromossômicas não histona** também é encontrada nos cromossomos eucarióticos. De vez em quando, histonas variantes, com algumas sequências diferentes de aminoácidos, são incorporadas na cromatina no lugar de uma das principais histonas. Resolva o Problema 18

Quadro 11.1 Características da eucromatina e da heterocromatina.

Característica	Eucromatina	Heterocromatina
Condensação da cromatina	Menos condensada	Mais condensada
Localização	Nos braços do cromossomo	Nos centrômeros, telômeros e em outros locais específicos
Tipo de sequências	Sequências únicas	Sequências repetidas*

Existência de genes	Muitos genes	Alguns genes*
Quando replicada	Durante toda a fase S	Fase S tardia
Transcrição	Frequente	Rara
Crossing over	Comum	Incomum

*Aplica-se apenas à heterocromatina constitutiva.

Conceitos

A cromatina, composta por DNA complexado a proteínas, é o material que constitui os cromossomos eucarióticos. As mais abundantes dessas proteínas são os cinco tipos de proteínas histona com carga elétrica positiva: H1, H2A, H2B, H3 e H4. Às vezes, histonas diferentes podem ser incorporadas na cromatina no lugar das histonas normais.

Checagem dos conceitos 3

Qual seria o efeito da neutralização das cargas positivas nas proteínas histona?

a. Elas se ligariam firmemente ao DNA.
b. Elas se ligariam menos firmemente ao DNA.
c. Elas não estariam mais atraídas uma pela outra.
d. Elas levariam ao superenrolamento do DNA.

Nucleossomo. A cromatina tem uma estrutura muito complexa com vários níveis de organização (**Figura 11.3**). O nível mais simples é a estrutura de dupla-hélice do DNA, discutida no Capítulo 10. Em um nível mais complexo, a molécula de DNA está associada a proteínas e está muito dobrada para produzir um cromossomo.

Quando a cromatina é isolada do núcleo de uma célula e visualizada com um microscópio eletrônico, ela aparece como um colar de pérolas (**Figura 11.4 A**). Se uma pequena quantidade de nuclease é adicionada a essa estrutura, a enzima rompe o "colar" entre as "pérolas", deixando as pérolas individuais presas a cerca de 200 pb de DNA (**Figura 11.4 B**). Se mais nuclease é adicionada, a enzima digere todo o DNA entre as pérolas e deixa um cerne de proteínas preso a um fragmento de DNA (**Figura 11.4 C**). Tais experimentos demonstraram que a cromatina não é uma associação aleatória de proteínas e DNA, mas tem uma estrutura fundamental repetida.

O cerne repetido de proteína e DNA produzido pela digestão com as enzimas de nuclease é o nível mais simples da estrutura da cromatina, o **nucleossomo** (ver Figura 11.3). O nucleossomo é uma partícula do cerne composta por DNA dobrado cerca de duas vezes ao redor de um octâmero de oito proteínas histona (duas cópias de H2A, H2B, H3 e H4), muito semelhante a um fio enrolado ao redor de um carretel (**Figura 11.4 D**). O DNA em contato direto com o octâmero de histonas tem entre 145 e 147 pb de comprimento.

Cada uma das proteínas histona que compõe a partícula cerne do nucleossomo tem uma "cauda" flexível, contendo de 11 a 37 aminoácidos, que se estende para fora do nucleossomo. Os aminoácidos com carga positiva encontrados nas caudas das histonas interagem com as cargas negativas dos fosfatos no DNA, mantendo o DNA e as histonas intimamente associadas. As caudas de um nucleossomo também podem interagir com os nucleossomos vizinhos, o que facilita a compactação destes. As modificações químicas nas caudas de histona provocam mudanças na estrutura da cromatina (discutida na próxima seção), necessárias à expressão do gene.

O quinto tipo de histona, H1, não é uma parte da partícula do cerne do nucleossomo, mas tem um papel importante na estrutura do nucleossomo. A H1 liga de 20 a 22 pb de DNA onde o DNA se une e deixa o octâmero (ver Figura 11.3), ajudando a manter o DNA no lugar, como um grampo ao redor do octâmero do nucleossomo.

Cada nucleossomo abriga cerca de 167 pb de DNA. Os nucleossomos estão localizados em intervalos regulares ao longo da molécula de DNA e estão separados um do outro pelo **DNA ligador** (*linker DNA*), que tem tamanho variado entre os tipos de células; na maioria delas, o DNA ligador inclui entre 30 e 40 pb. As proteínas cromossômicas não histona podem estar associadas a esse DNA ligador, e algumas também parecem se ligar diretamente à partícula do cerne. **Resolva os Problemas 19 e 22**

Estrutura da cromatina de ordem superior. Quando a cromatina está em forma condensada, os nucleossomos adjacentes não estão separados por espaço igual ao comprimento do DNA ligador; em vez disso, os nucleossomos dobram

sobre si mesmos para formar uma estrutura densa, altamente condensada (ver Figura 11.3) que compõe uma fibra com diâmetro de 30 nm. A estrutura molecular exata da fibra de 30 nm ainda é desconhecida.

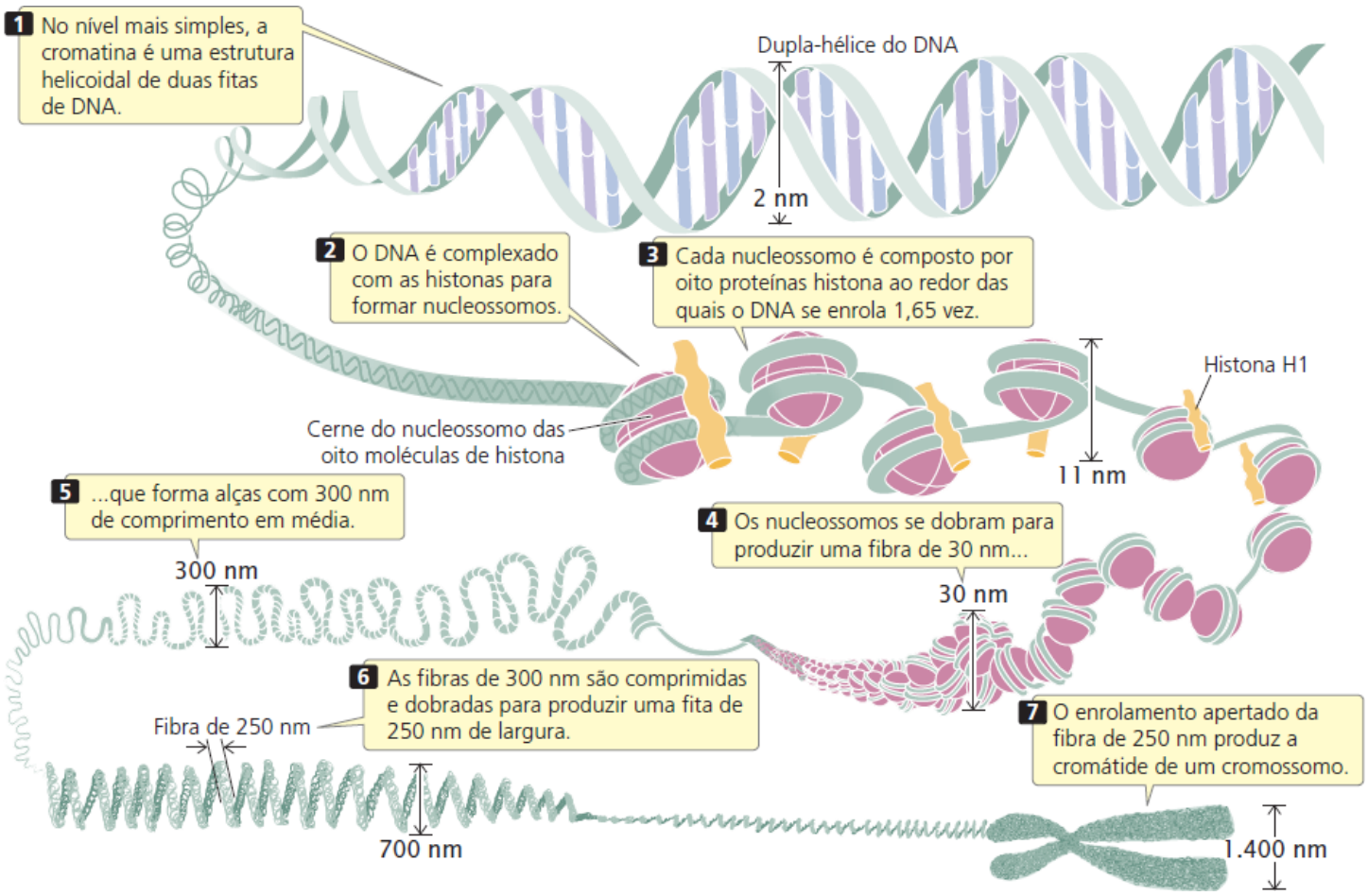


Figura 11.3 A cromatina tem uma estrutura muito complexa com vários níveis de organização.

O próximo nível superior da estrutura da cromatina é uma série de alças de fibras de 30 nm (Figura 11.3), cada uma ancorada em sua base por proteínas. Em média, cada alça inclui 20.000 a 100.000 pb de DNA e mede cerca de 300 nm de comprimento, mas as alças individuais variam muito. As fibras de 300 nm são comprimidas e dobradas para produzir uma fita de 250 nm de largura. O rígido enrolamento helicoidal da fibra de 250 nm, por sua vez, produz a estrutura que aparece na metáfase – cromátides individuais de aproximadamente 700 nm de comprimento.

Conceitos

O nucleossomo consiste em uma cerne de oito proteínas histona e DNA que envolve o cerne. Uma única histona H1 se associa com cada partícula do cerne. Os nucleossomos são separados por DNA ligador. Os nucleossomos se dobram para formar uma fibra de cromatina de 30 nm, que aparece como várias alças condensadas que criam uma fibra de 250 nm. O enrolamento helicoidal da fibra de 250 nm produz uma cromátide.

Checagem dos conceitos 4

Quantas cópias da histona H2B seriam encontradas na cromatina contendo 50 nucleossomos?

a. 5.
b. 10.
c. 50.
d. 100.

Mudanças na estrutura da cromatina

Embora o DNA eucariótico precise estar bem condensado para caber no núcleo da célula, ele também precisa ser desenrolado

periodicamente para que ocorra transcrição e replicação.

Cromossomos politênicos. Cromossomos gigantes chamados de **cromossomos politênicos** são encontrados em alguns tecidos da *Drosophila* e outros organismos (**Figura 11.5**). Os cromossomos politênicos forneceram indícios da natureza transformante da estrutura da cromatina. Esses grandes cromossomos incomuns surgem quando ciclos repetidos de replicação de DNA ocorrem sem divisões celulares, produzindo milhares de cópias de DNA que ficam lado a lado. Várias bandas são reveladas quando esses cromossomos são destacados com corantes. Sob certas condições, as bandas exibem ***puffs* cromossômicos** – intumescências localizadas do cromossomo. Cada *puff* é uma região da cromatina que tem uma estrutura relaxada e, consequentemente, um estado mais aberto. A pesquisa indica que os *puffs* cromossômicos são regiões de transcrição ativa. Essa correlação entre a ocorrência da transcrição e o relaxamento da cromatina no sítio do *puff* indica que a estrutura da cromatina sofre mudança dinâmica associada à atividade do gene. **Resolva o Problema 20**

Sensibilidade à enzima DNase I. Uma segunda evidência de mudança da estrutura da cromatina com a atividade dos genes é a sensibilidade à DNase I, uma enzima que digere DNA. A capacidade dessa enzima em digerir DNA depende da estrutura da cromatina: quando o DNA está firmemente preso a proteínas histona, é menos sensível à DNase I, enquanto o DNA não ligado é mais sensível à DNase I. Os resultados dos experimentos que examinam o efeito da DNase I em genes específicos mostram que a sensibilidade à DNase está correlacionada com a atividade dos genes. Estudos dos genes da globina da galinha evidenciam essa correlação. Os genes da globina codificam a hemoglobina nos eritroblastos (precursores das hemácias) das galinhas (**Figura 11.6**). Esses tipos de experimentos demonstram que os genes ativos para transcrição são sensíveis à DNase I, indicando que a estrutura da cromatina está mais exposta durante a transcrição.

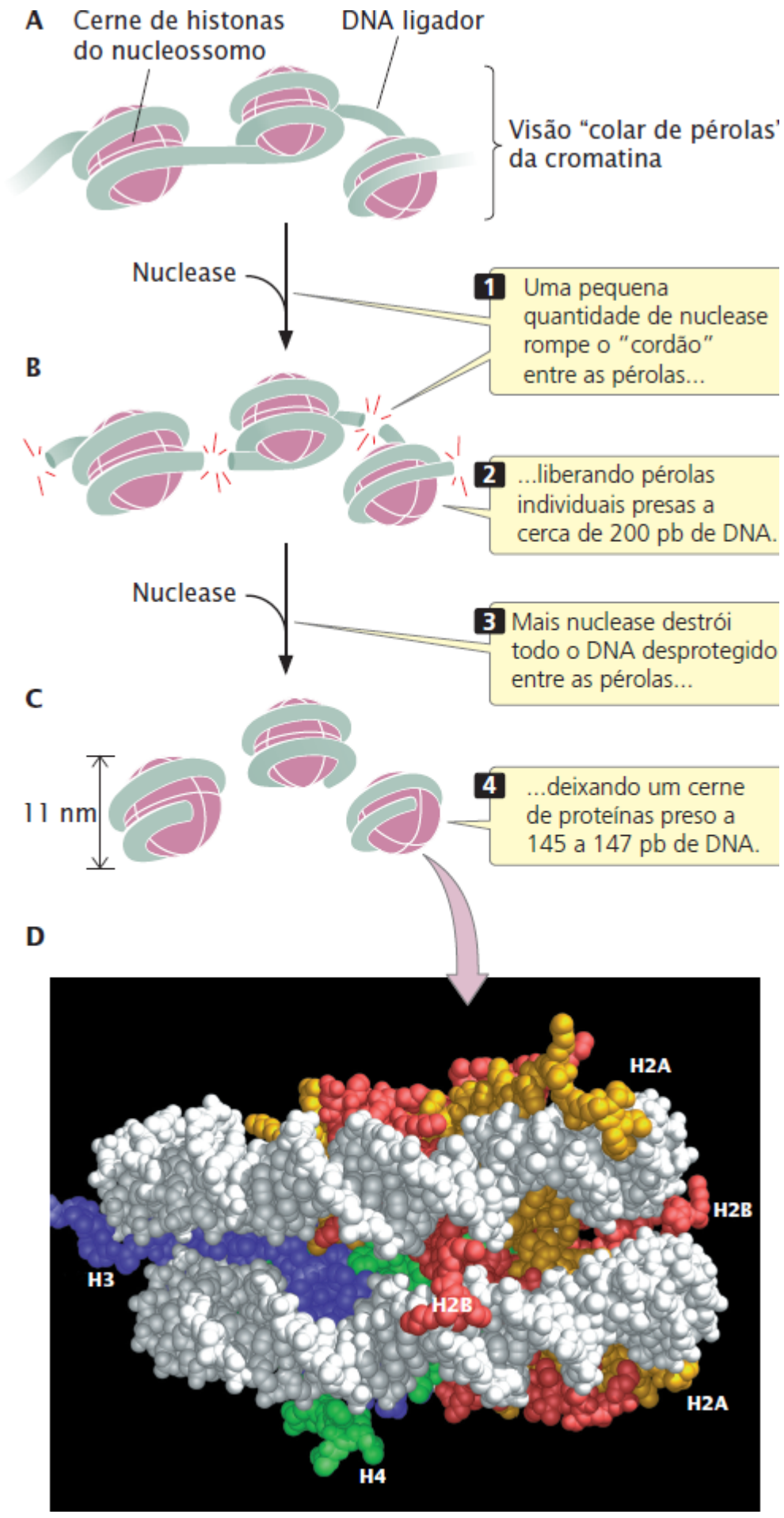


Figura 11.4 O nucleossomo é a unidade repetida fundamental da cromatina. A parte D mostra um modelo de preenchimento espacial da partícula do cerne, que consiste em duas cópias de cada H2A, H2B, H3 e H4 ao redor dos quais o DNA (*branco*) se enrola. (*Parte D*: Reimpressa com permissão de Macmillan Publishers Ltd. De K. Luger et al., Nature 389:251. © 1997. Cortesia de T. H. Richmond.)

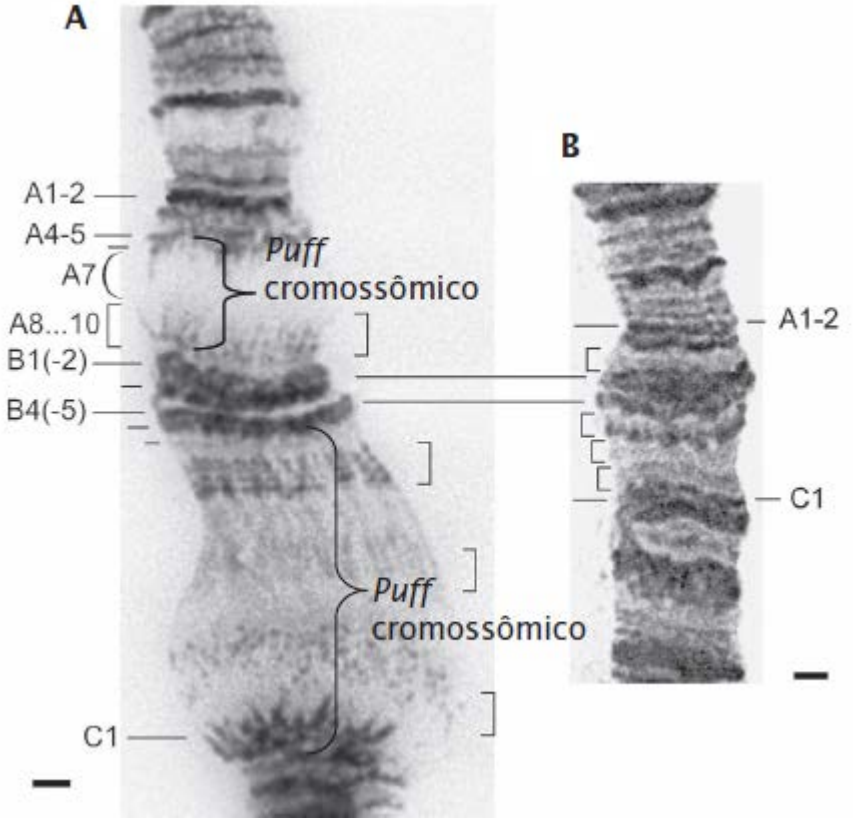


Figura 11.5 Os *puffs* cromossômicos são regiões de cromatina relaxada onde ocorre a transcrição ativa. Indicados aqui estão (**A**) *puffs* cromossômicos de cromossomos politênicos gigantes isolados das glândulas salivares da larva da *Drosophila* e (**B**) a região correspondente sem *puffs* cromossômicos. (Cortesia de Dmitri V. Novikov.)

Qual é a natureza da alteração da estrutura da cromatina que produz *puffs* cromossômicos e a sensibilidade à DNase I? Em ambos os casos, a cromatina relaxa; aparentemente, as histonas afrouxam sua fixação ao DNA. Um processo que altera a estrutura da cromatina é a acetilação. Enzimas chamadas acetiltransferases prendem grupos acetila aos aminoácidos lisina nas caudas da histona. Essa modificação reduz as cargas elétricas positivas que normalmente existem na lisina e desestabiliza a estrutura do nucleossomo, e então as histonas seguram o DNA com menos firmeza. Outras modificações químicas das proteínas histona, como metilação e fosforilação, também alteram a estrutura da cromatina, como as proteínas especiais de remodelagem da cromatina que se ligam ao DNA.

Alterações epigenéticas associadas a modificações da cromatina. Vamos observar como a estrutura da cromatina pode ser alterada por meio de modificação química das proteínas histona. Várias outras mudanças também podem afetar a estrutura da cromatina, incluindo a metilação do DNA (ver *Capítulo 10*), o uso de histonas diferentes no nucleossomo e a ligação de proteínas ao DNA e à cromatina. Embora essas mudanças não alterem a sequência de DNA, elas têm efeitos importantes na expressão dos genes, assunto que será discutido com mais detalhes no Capítulo 17.

Algumas mudanças na estrutura da cromatina são retidas por meio da divisão celular, e então elas são transmitidas para as gerações futuras das células e até para gerações futuras dos organismos em alguns casos. Alterações estáveis da estrutura da cromatina que podem ser passadas para células ou organismos individuais são chamadas de **alterações epigenéticas**, ou simplesmente epigenética (ver *Capítulo 5*). Por exemplo, o *locus agouti* ajuda a determinar a cor do pelo dos camundongos: os genitores com sequências de DNA idênticas, mas graus diferentes de metilação em seu DNA, podem dar origem a descendentes com diferentes cores de pelo. Tais alterações epigenéticas foram observadas em vários organismos e são responsáveis por diversos efeitos fenotípicos. Diferente das mutações, as alterações epigenéticas não alteram a sequência de DNA, podem ser revertidas e muitas vezes são influenciadas por fatores ambientais.

Experimento

Pergunta: A estrutura da cromatina é alterada na transcrição?

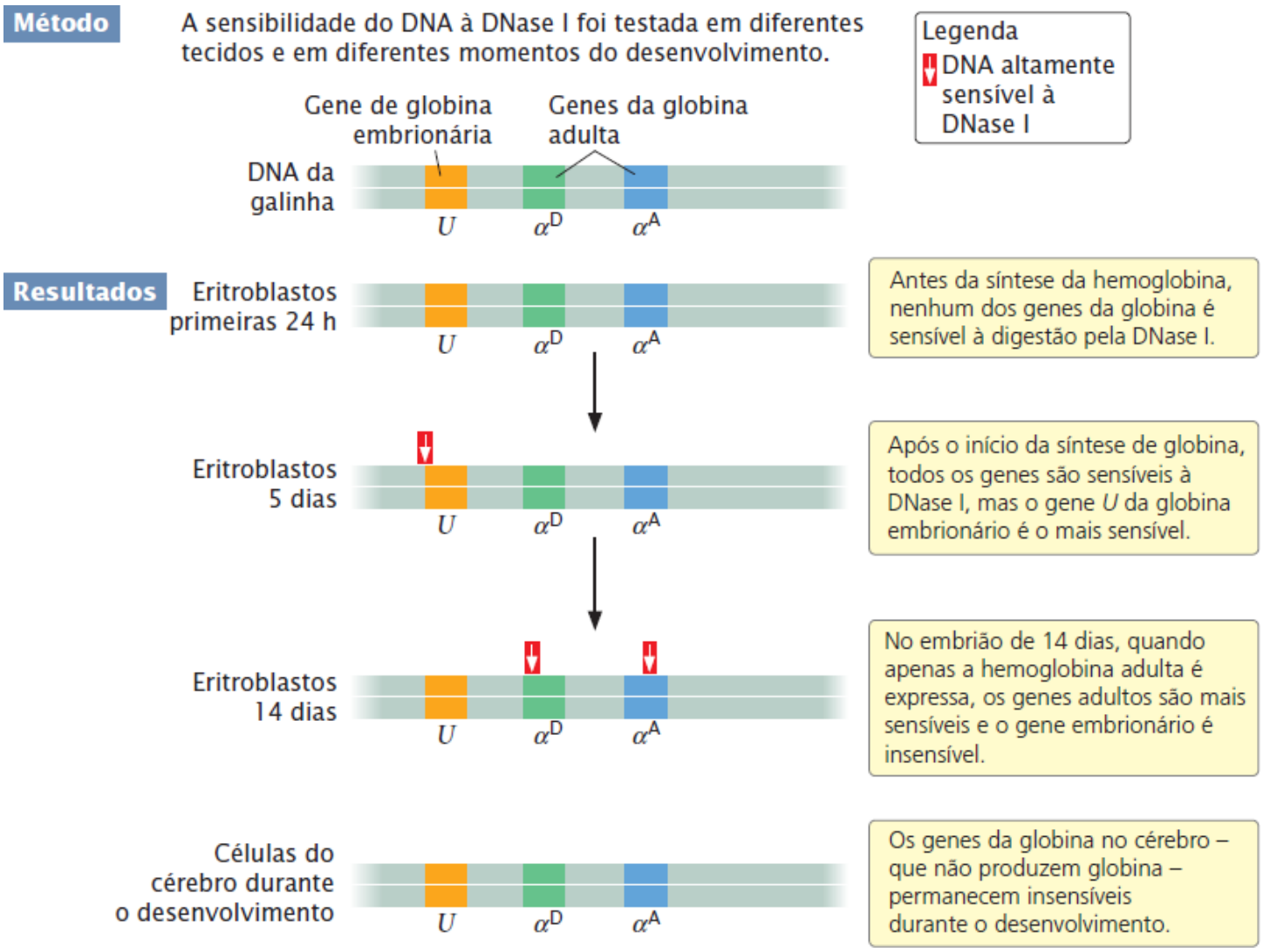


Figura 11.6 A sensibilidade à enzima DNase I está correlacionada com a transcrição dos genes da globina nos eritroblastos dos embriões de galinha. O gene *U* codifica a hemoglobina embrionária; os genes α^D e α^A codificam a hemoglobina adulta.

Conceitos

As alterações epigenéticas são modificações da cromatina ou da estrutura do DNA que não incluem mudanças na sequência de bases, mas são estáveis e transmitidas para células ou organismos. Algumas alterações epigenéticas são resultado de alterações das histonas.

11.2 Os cromossomos eucarióticos têm centrômeros e telômeros

Os cromossomos se separam na mitose e meiose e permanecem estáveis durante muitas divisões celulares. Essas propriedades dos cromossomos surgem a partir das características estruturais especiais dos cromossomos, incluindo os centrômeros e telômeros.

Estrutura do centrômero

O centrômero é uma região contraída do cromossomo ao qual as fibras de fuso se prendem e é essencial para o movimento correto do cromossomo na mitose e meiose (ver *Capítulo 2*). O papel essencial do centrômero no movimento do cromossomo foi identificado pelos primeiros geneticistas, que observaram as consequências da ruptura do cromossomo. Quando uma ruptura de cromossomo produz dois fragmentos, um com um centrômero e um sem (**Figura 11.7**), o fragmento de cromossomo com centrômero se fixa às fibras do fuso e se move na direção do polo. O fragmento sem centrômero não se conecta à fibra do fuso e é perdido porque não consegue se mover para o núcleo da célula irmã durante a mitose (ver Figura 11.7).

Os centrômeros são sítios de ligação para o cinetócoro, ao qual as fibras do fuso se ligam. Na *Drosophila*, na *Arabidopsis* e nos seres humanos, os centrômeros transportam centenas de milhares de pares de bases. A maior parte do centrômero é feita de heterocromatina. Surpreendentemente, não existem sequências específicas encontradas em todos os centrômeros, o que levanta a questão sobre o que determina exatamente onde o centrômero está. A pesquisa sugere que a maioria dos centrômeros não está definida por sequência de DNA, mas por alterações epigenéticas na estrutura da cromatina. Os nucleossomos nos centrômeros da maioria dos eucariotos têm uma proteína histona diferente chamada CenH3, que toma o lugar da histona H3. A histona CenH3 traz uma mudança no nucleossomo e na estrutura da cromatina, que possivelmente promove a formação do cinetócoro e a fixação das fibras do fuso ao cromossomo.

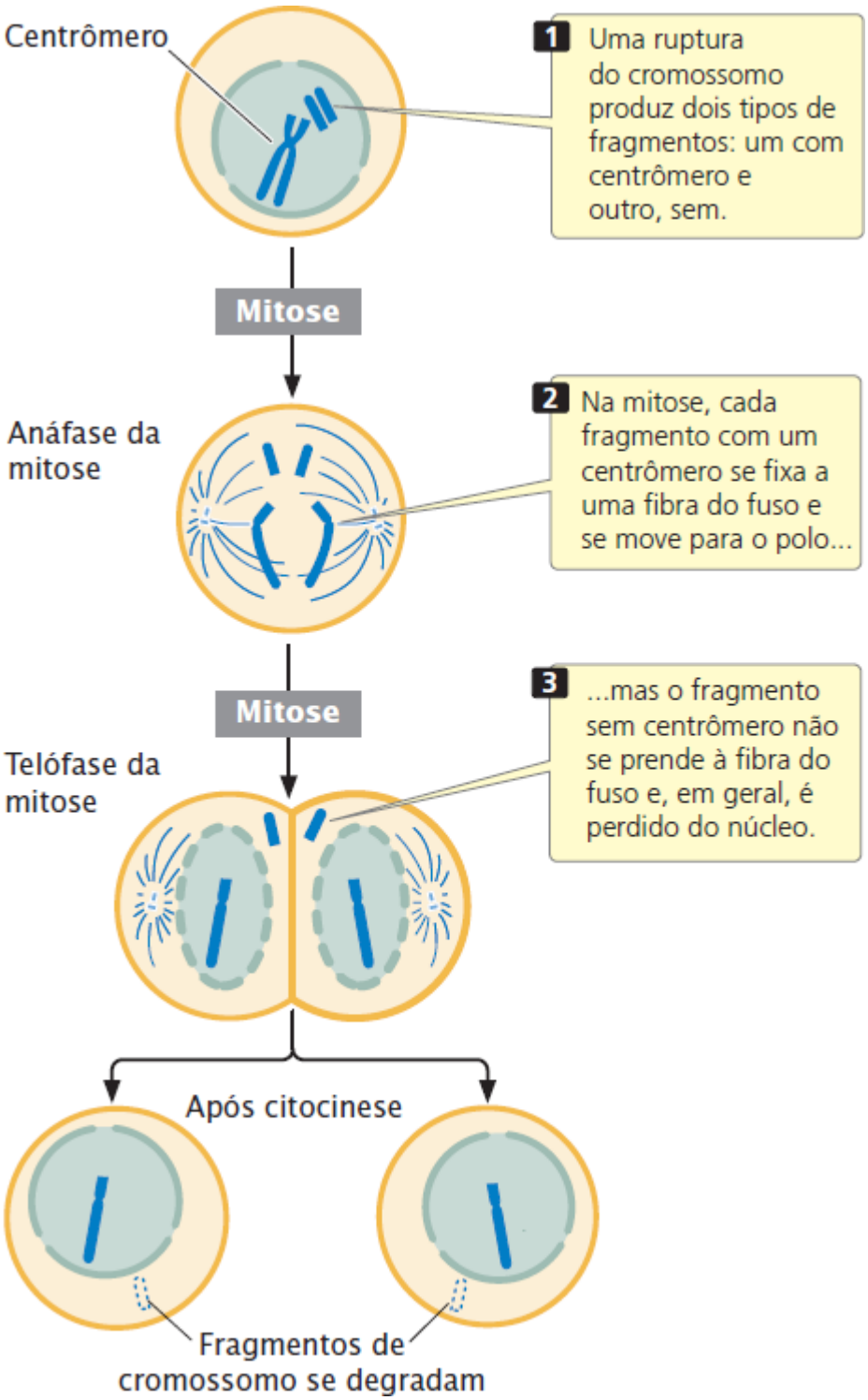


Figura 11.7 **Os fragmentos do cromossomo sem centrômeros são perdidos na mitose.**

Conceitos

O centrômero é uma região do cromossomo onde as fibras do fuso se prendem. Os centrômeros apresentam variação considerável na estrutura e são diferenciados por alterações epigenéticas na estrutura da cromatina,

incluindo o uso da histona H3 diferente no nucleossomo.

Checagem dos conceitos 5

O que acontece ao cromossomo que perde seu centrômero?

Estrutura do telômero

Os telômeros são as extremidades naturais de um cromossomo (ver Figura 2.6 e a introdução deste capítulo). O trabalho pioneiro de Hermann Muller (nas moscas-da-fruta) e Barbara McClintock (no milho) mostrou que as rupturas do cromossomo produzem extremidades instáveis com tendência a se unirem e capazes de degradar o cromossomo. Como a fixação e a degradação não ocorrem nas extremidades de um cromossomo com telômeros, cada telômero funciona obrigatoriamente como uma tampa que estabiliza o cromossomo. Os telômeros também fornecem meios para replicar as extremidades do cromossomo, assunto que será discutido no Capítulo 12. Em 2009, Elizabeth Blackburn, Carol Greider e Jack Szostak receberam o prêmio Nobel em Fisiologia e Medicina por descobrirem a estrutura dos telômeros e como eles são replicados (discutido no *Capítulo 12*).

Atualmente, os telômeros de protozoários, plantas, seres humanos e outros organismos foram isolados. Eles apresentam estrutura semelhante em sua maioria (**Quadro 11.2**). Essas **sequências teloméricas** são compostas por unidades repetidas de uma série de nucleotídios adenina ou timina seguida por vários nucleotídios guanina, assumindo a forma 5′-(A ou T)$_m$G$_n$-3′, em que *m* é de 1 a 4 e *n* é 2 ou mais. Por exemplo, a unidade repetida nos telômeros humanos é 5′-TTAGGG-3′, que pode ser repetida de centenas a milhares de vezes. A sequência é sempre orientada com o cordão de Gs e Cs na direção da extremidade do cromossomo, como mostrado aqui:

Em direção ao centrômero	←	5′-TTAGGGTTAGGGTTAGGG-3′ 3′-AATCCCAATCCCAATCCC-5′	→	Extremi-dade do cromos-somo

A fita rica em G se sobressai do cordão complementar rico em C na extremidade do cromossomo (**Figura 11.8 A**) e é chamado de protuberância 3′. A protuberância 3′ nos telômeros dos mamíferos mede de 50 a 500 nucleotídios. Proteínas especiais se ligam à sequência de fita única rica em G, protegendo o telômero da degradação e evitando que as extremidades dos cromossomos grudem uma na outra. Um complexo de multiproteínas chamado de **shelterina** se liga aos telômeros e protege as extremidades do DNA de serem inadvertidamente consertadas como uma ruptura da fita dupla no DNA. Em algumas células, a protuberância de fita única pode se dobrar e parear com um curto trecho de DNA para formar uma estrutura chamada de laço em t, que também funciona na proteção da extremidade do telômero, evitando sua degradação (**Figura 11.8 B**).

Quadro 11.2 Sequências de DNA encontradas nos telômeros de vários organismos.

Organismo	Sequência
Tetrahymena (protozoário)	5′-T T GGGG-3′ 3′-AAC C C C-5′
Saccharomyces (levedura)	5′-T$_{1-6}$GTG$_{2-3}$-3′ 3′-A$_{1-6}$CAC$_{2-3}$-5′
Caenorhabditis (nematódeo)	5′-T TAGGC-3′ 3′-AATC C G-5′
Vertebrado	5′-T TAGGG-3′ 3′-AAT CC C-5′
Arabidopsis (planta)	5′-T T TAGGG-3′ 3′-AAAT CCC-5′

Fonte: V. A. Zakian, Science 270:1602, 1995.

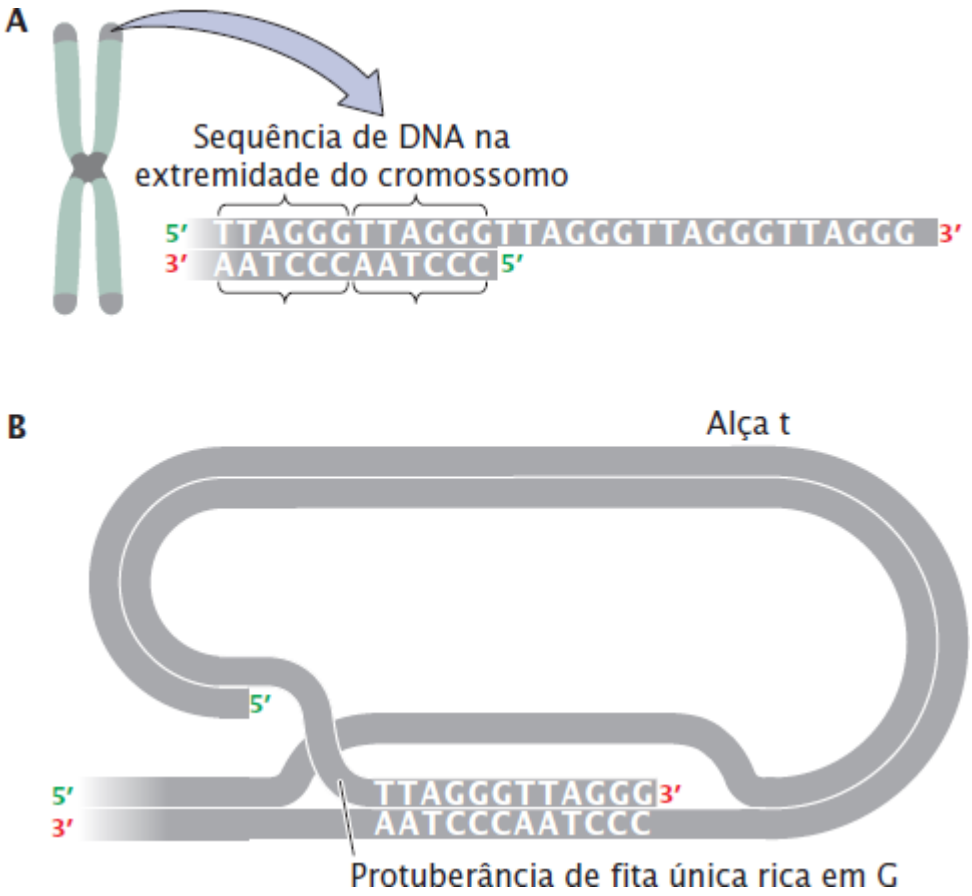


Figura 11.8 O DNA nas extremidades dos cromossomos eucarióticos é composto por sequências teloméricas. A. A fita rica em G no telômero é mais longa que a fita rica em C. **B.** Em algumas células, a fita rica em G se dobra e pareia com uma tira curta do DNA para formar uma alça t.

Conceitos

Um telômero é a extremidade estabilizadora de um cromossomo. Na extremidade de cada telômero há várias sequências teloméricas curtas.

Checagem dos conceitos 6

Qual é a característica das sequências de DNA nos telômeros?

a. Uma fita é composta pelos nucleotídios guanina e adenina (ou timina).
b. São compostas por sequências repetidas.
c. Uma fita se sobressai sobre a outra, criando um DNA de fita única na extremidade.
d. Todas as opções acima.

11.3 O DNA eucariótico contém as várias classes de variação de sequências

Os organismos eucarióticos apresentam diferenças drásticas na quantidade de DNA por célula, chamada de **valor de C** do organismo (**Quadro 11.3**). Cada célula de uma mosca-da-fruta, por exemplo, tem 35 vezes a quantidade de DNA encontrada em uma célula da *E. coli*. Em geral, as células eucarióticas têm mais DNA que as procarióticas, mas a variação nos valores de C de diferentes eucariotos é imensa. As células humanas têm cerca de 10 vezes a quantidade de DNA encontrada nas células da *Drosophila*, enquanto as células de algumas salamandras têm 20 vezes a quantidade encontrada nas células humanas. Claramente, essas diferenças no valor de C não podem ser explicadas apenas pelas diferenças na complexidade do organismo. Então, o que esse DNA extra está fazendo nas células eucarióticas? Essa questão é chamada de **paradoxo do valor de C**. Não temos uma resposta definitiva para esse paradoxo, mas as sequências de DNA eucariótico revelam uma complexidade que não existe no DNA procariótico.

Quadro 11.3 Tamanho dos genomas de vários organismos.

Organismo	Tamanho aproximado do genoma (pb)
λ (bacteriófago)	50.000
Escherichia coli (bactéria)	4.640.000
Saccharomyces cerevisiae (levedura)	12.000.000
Arabidopsis thaliana (planta)	125.000.000
Drosophila melanogaster (inseto)	170.000.000
Homo sapiens (humano)	3.200.000.000
Zea mays (milho)	4.500.000.000
Amphiuma (salamandra)	765.000.000.000

Desnaturação e renaturação do DNA

O primeiro indício de que o DNA eucariótico tem vários tipos de sequências não encontradas no DNA procariótico veio de estudos em que o DNA de fita dupla foi separado e reassociado. Quando o DNA de fita dupla em solução é aquecido, as pontes de hidrogênio que mantêm as duas fitas unidas são enfraquecidas e, com calor suficiente, as duas fitas de nucleotídios se separam completamente, um processo chamado de **desnaturação**. A temperatura na qual o DNA desnatura, chamada **temperatura de desnaturação** (T_m) depende da sequência de bases da amostra particular de DNA: os pares de base G-C têm três pontes de hidrogênio, enquanto os pares A-T têm apenas duas; então, a separação dos pares G-C requer mais calor (energia) do que a dos pares A-T.

A desnaturação do DNA pelo calor é reversível; se o DNA de fita única for lentamente resfriado, as fitas únicas colidem e as pontes de hidrogênio serão formadas novamente entre os pares de base complementares, produzindo o DNA de fita dupla. Essa reação é chamada de **renaturação** ou reanelamento.

Duas moléculas de fita única de DNA de diferentes fontes, com diferentes organismos, vão anelar se forem complementares, um processo chamado de **hibridização**. Para que ocorra a hibridização, as duas fitas não são complementares em todas as suas bases – apenas o suficiente para manter as duas fitas juntas. A extensão da hibridização pode ser usada para medir a semelhança dos ácidos nucleicos a partir de duas fontes diferentes e para avaliar as relações evolutivas. A taxa na qual a hibridização ocorre também fornece informações sobre a complexidade da sequência de DNA. Resolva o Problema 26

Tipos de sequências de DNA nos eucariotos

O DNA eucariótico é composto por pelo menos três tipos de sequências: DNA de sequência única, DNA moderadamente repetitivo e DNA altamente repetitivo. O **DNA de sequência única** é composto por sequências que ocorrem apenas uma vez ou, no máximo, algumas vezes no genoma. Esse DNA inclui sequências que codificam proteínas, assim como uma enorme quantidade de DNA cuja função permanece desconhecida. Os genes que se encontram em uma única cópia constituem grosseiramente 25 a 50% dos genes que codificam proteínas na maioria dos eucariotos multicelulares. Outros genes com DNA de sequência única ocorrem em cópias muito semelhantes, mas não idênticas, e são chamados de **família de genes**. A maioria das famílias de genes surge da duplicação de um gene existente e inclui alguns genes como membros, mas outras, como as que codificam as proteínas imunoglobulinas nos vertebrados, têm centenas de membros. Os genes que codificam as globinas semelhantes à β-globina são outro exemplo de família de genes. Nos seres humanos, há sete genes da β-globina, agrupados no cromossomo 11. Os polipeptídios codificados por esses genes se unem a polipeptídios da α-globina para formar as moléculas de hemoglobina, que transportam oxigênio no sangue.

Outras sequências existem em muitas cópias e são chamadas de **DNA repetitivo**. Alguns organismos eucarióticos têm grandes quantidades de DNA repetitivo; por exemplo, quase metade do genoma humano é composto por esse tipo. Uma classe principal de DNA repetitivo é chamada de **DNA moderadamente repetitivo**, composto por sequências de 150 a 300 pb (embora possam ser maiores) que são repetidas milhares de vezes. Algumas dessas sequências executam funções importantes para a célula; por exemplo, os genes para RNAs ribossômicos (rRNAs) e RNAs de transferência (tRNAs) compõem uma parte do DNA moderadamente repetitivo. O próprio DNA moderadamente repetitivo é composto por dois tipos de repetições. As **sequências de repetição em *tandem*** aparecem uma após a outra e tendem a agrupadar-se em localizações particulares nos cromossomos. As **sequências repetidas intercaladas** estão dispersas por todo o genoma. Um exemplo de repetição intercalada é a sequência *Alu*, de aproximadamente 300 pb, encontrada em mais de um milhão de vezes e que engloba 11% do genoma

humano, embora não tenha função celular clara. Repetições curtas, como as sequências *Alu*, são chamadas de **SINEs (elementos curtos intercalados)**. As repetições intercaladas mais longas, compostas por vários milhares de pares de bases, são chamadas de **LINEs (elementos longos intercalados)**. Uma classe de LINE, chamada LINE1, inclui cerca de 17% do genoma humano. A maioria das repetições intercaladas são as remanescentes de elementos de transposição, sequências que podem ser multiplicar e mover (ver *Capítulo 18*).

A outra principal classe de DNA repetitivo é o **DNA altamente repetitivo**. Essas sequências curtas, em geral com 10 pb de comprimento, são encontradas em centenas de milhares a milhões de cópias que estão repetidas em *tandem* e agrupadas em algumas regiões do cromossomo, em especial nos centrômeros e telômeros. O DNA altamente repetitivo também é chamado de DNA satélite, porque sua porcentagem das quatros bases é diferente da porcentagem de outras sequências de DNA e, portanto, ele se separa como uma fração satélite quando centrifugado em altas velocidades em um gradiente de densidade (ver *Capítulo 12*). É raro o DNA altamente repetitivo ser transcrito em RNA. Embora essas sequências possam contribuir para a função do centrômero e do telômero, a maior parte do DNA repetitivo não tem função conhecida.

As reações de renaturação do DNA e, mais recentemente, o sequenciamento direto de genomas eucarióticos nos diz muito sobre como a informação genética é organizada dentro dos cromossomos. Agora sabemos que a densidade dos genes varia muito entre e dentro dos cromossomos. Por exemplo, o cromossomo 19 humano tem alta densidade de genes, com cerca de 26 genes por milhões de pares de base. O cromossomo 13, por outro lado, tem apenas 6,5 genes por milhão de pares de base. A densidade do gene também pode variar em regiões diferentes do mesmo cromossomo: algumas partes do braço longo do cromossomo 13 têm apenas 3 genes por milhão de pares de base, enquanto outras têm quase 30 genes por milhão de pares de base. E o braço curto do cromossomo 13 quase não tem gene, composto praticamente por heterocromatina.

O papel funcional das sequências de DNA que não codificam proteínas, incluindo o DNA repetitivo, tornou-se recentemente assunto de estudo pelo projeto de Enciclopédia de Elementos do DNA (ENCODE) (ver *Capítulo 20*). O objetivo do ENCODE era identificar todos os nucleotídios no genoma humano com alguma função. O projeto concluiu que boa parte do genoma é transcrito e pelo menos 80% das sequências são funcionais. Muitas das sequências funcionais parecem ajudar a controlar a expressão do gene.

Conceitos

O DNA eucariótico compreende três classes importantes: DNA de sequência única, DNA moderadamente repetitivo e DNA altamente repetitivo. O DNA de sequência única é composto por sequências que existem em uma ou algumas cópias; o DNA moderadamente repetitivo é composto por sequências que podem ser muitas centenas de pares de bases de comprimento e encontra-se em milhares ou centenas de milhares de cópias. O DNA altamente repetitivo é composto por sequências muito curtas repetidas em *tandem* e encontra-se em centenas de milhares a milhões de cópias. A densidade dos genes varia muito entre e até dentro dos cromossomos.

Checagem dos conceitos 7

A maioria dos genes que codificam as proteínas é encontrada:

a. No DNA de sequência única.
b. No DNA moderadamente repetitivo.
c. No DNA altamente repetitivo.
d. Todas as opções acima.

11.4 O DNA de organelas tem características únicas

Como observamos, os cromossomos eucarióticos estão localizados no núcleo e têm uma estrutura complexa composta por DNA e histonas associadas. Entretanto, parte do DNA encontrado nas células eucarióticas está fora do núcleo, tem uma organização diferente e exibe um padrão diferente de herança em relação ao DNA nuclear. Esse DNA é encontrado nas mitocôndrias e nos cloroplastos, que são organelas ligadas a membranas localizadas no citoplasma das células eucarióticas (**Figura 11.9**).

Estrutura da mitocôndria e do cloroplasto

As mitocôndrias são encontradas em quase todas as células eucarióticas, enquanto os cloroplastos são encontrados em plantas e em alguns protistas. Ambas as organelas geram ATP, o transportador universal de energia das células.

As mitocôndrias são estruturas tubulares com 0,5 a 1,0 micrometro (μm) de diâmetro, próximo do tamanho de uma típica

bactéria, enquanto os cloroplastos têm tipicamente cerca de 4 a 6 μm de diâmetro. Ambos são circundados por duas membranas que englobam uma região (chamada de matriz nas mitocôndrias e de estroma nos cloroplastos) que contém enzimas, ribossomos, RNA e DNA. Nas mitocôndrias, a membrana interna está muito dobrada; dentro dela estão incrustadas enzimas que catalisam o transporte de elétrons e a fosforilação oxidativa. Os cloroplastos têm uma membrana tilacoide, que está muito dobrada e empilhada para formar agregados chamados de grana. Essa membrana tem os pigmentos e enzimas necessários para a fotofosforilação. Novas mitocôndrias e cloroplastos surgem a partir da divisão das organelas existentes – divisões que ocorrem durante todo o ciclo celular e são independentes da mitose e meiose.

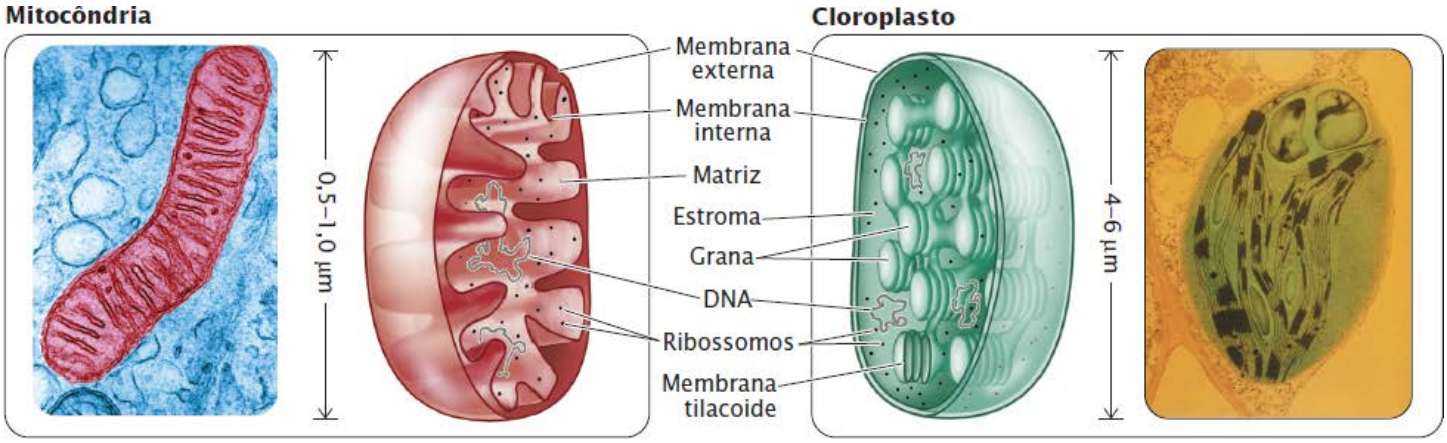


Figura 11.9 Comparação das estruturas das mitocôndrias e cloroplastos. (*Esquerda*: Don W. Fawcett/Science Source/Photo Researchers, Inc. *Direita*: Biophoto Associates/Photo Researchers.)

As mitocôndrias e os cloroplastos têm DNA que codifica alguns polipeptídios usados pela organela, assim como RNA encontrado no ribossomo (RNA ribossômico ou rRNA) e alguns RNA de transferência (tRNAs) necessários para a tradução dessas proteínas. Os genes para a maioria das 900 ou mais proteínas estruturais e enzimas encontradas nas mitocôndrias são codificados pelo DNA nuclear; o genoma mitocondrial codifica apenas algumas proteínas e algumas moléculas de rRNA e tRNA necessárias para a síntese de proteínas mitocondriais.

Teoria endossimbiótica

Os cloroplastos e as mitocôndrias têm muitas semelhanças com as bactérias. Essa semelhança não é superficial; de fato, há evidências convincentes de que essas organelas evoluíram a partir das eubactérias (ver *Capítulo 2*). A **teoria endossimbiótica** (**Figura 11.10**) propõe que mitocôndrias e cloroplastos eram bactérias livres que se tornaram habitantes internos (endossimbiontes) das primeiras células eucarióticas. Presume-se que, com a evolução, a maior parte dos genes originais dos endossimbiontes foi perdida (porque os genes nucleares existentes forneciam a mesma função) ou foi transferida para o núcleo.

Muitas evidências apoiam a ideia de que mitocôndrias e cloroplastos se originaram como células eubacterianas. Muitos eucariotos unicelulares modernos (protistas) são hospedeiros de bactérias endossimbióticas. As mitocôndrias e os cloroplastos têm o mesmo tamanho das atuais eubactérias e apresentam seu próprio DNA, com muitas características em comum com o DNA eubacteriano. As mitocôndrias e os cloroplastos têm ribossomos, e alguns apresentam tamanho e estrutura semelhantes aos dos ribossomos eubacterianos. Além disso, os antibióticos que inibem a síntese proteica nas eubactérias, mas não afetam a síntese nas células eucarióticas, também inibem a mesma síntese nessas organelas.

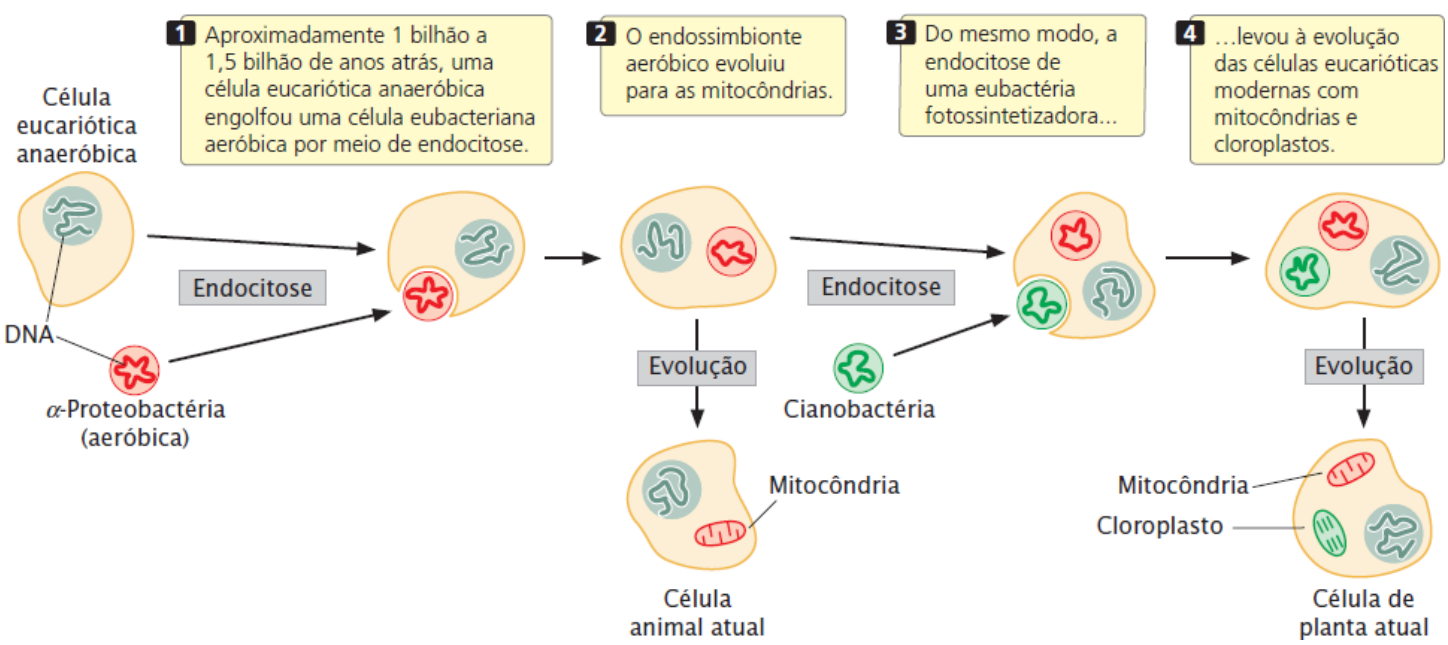


Figura 11.10 A teoria endossimbiótica propõe que as mitocôndrias e os cloroplastos nas células eucarióticas são oriundos de eubactérias.

A evidência mais forte para a teoria endossimbiótica surge com o estudo das sequências de DNA, que demonstra que as sequências no mtDNA e cpDNA estão mais próximas das sequências encontradas nos genes das eubactérias do que as encontradas no núcleo eucariótico. Todas essas evidências indicam que as mitocôndrias e os cloroplastos estão mais próximos das células eubacterianas do que das células eucarióticas onde são encontrados.

Conceitos

Mitocôndrias e cloroplastos são organelas envoltas pela membrana das células eucarióticas que apresentam seu próprio DNA. A teoria endossimbiótica propõe que essas organelas começaram como eubactérias livres e desenvolveram relações endossimbióticas estáveis com as primeiras células eucarióticas.

Checagem dos conceitos 8

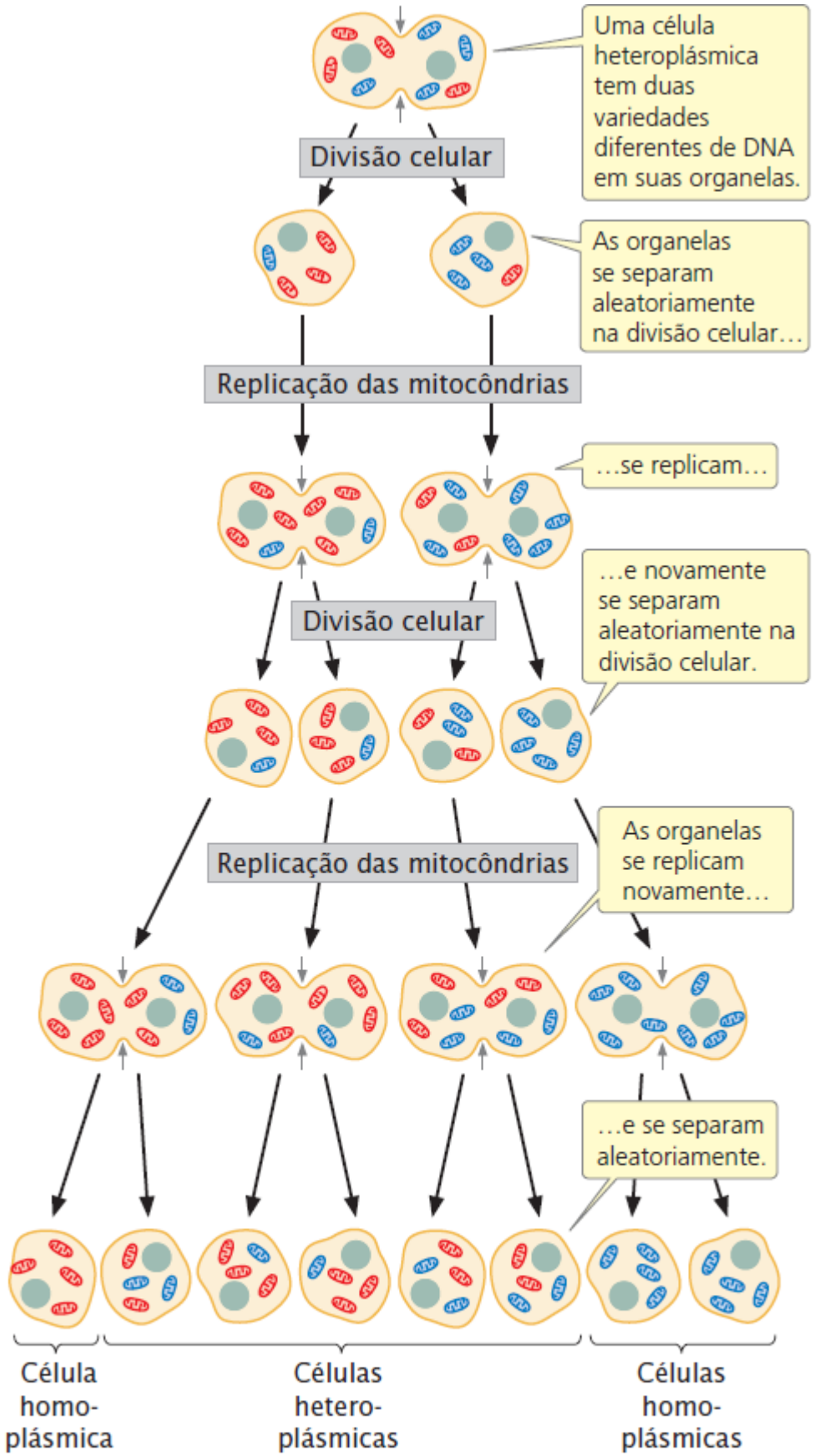
Que evidência apoia a teoria endossimbiótica?

Herança uniparental de traços codificados pelas organelas

As mitocôndrias e os cloroplastos são encontrados no citoplasma, como já foi afirmado aqui, e em geral são herdados de um único genitor. Assim, os traços codificados pelo DNA mitocondrial (mtDNA) e o DNA de cloroplasto (cpDNA) exibem herança uniparental (ver *Capítulo 5*). Nos animais, o mtDNA é herdado quase exclusivamente da genitora, embora tenha sido documentada a transmissão ocasional de mtDNA do genitor masculino. A herança materna do mtDNA animal pode ser em parte uma função do tamanho do gameta – os espermatozoides são muito menores do que os óvulos e têm poucas mitocôndrias. Entretanto, uma pesquisa recente descobriu que em alguns eucariotos, as mitocôndrias paternas são seletivamente eliminadas por autofagia, um processo no qual as mitocôndrias são digeridas pela célula. As mitocôndrias paternas são encaminhadas para destruição, enquanto as maternas, não. O mecanismo que produz essa diferença é desconhecido. A herança paterna das organelas é comum nos gimnospermas (coníferas) e em certas angiospermas (plantas florescentes). Algumas plantas até exibem herança biparental de mtDNA e cpDNA.

Segregação replicativa. As células individuais podem ter de dúzias a centenas de organelas, cada uma com várias cópias do genoma da organela, então cada célula tem de centenas a milhares de cópias de genomas de mitocôndria e cloroplasto. Uma mutação que surja em uma molécula de DNA da organela gera uma mistura de organelas dentro da célula, algumas com uma sequência de DNA mutante e outras com uma sequência de DNA do tipo selvagem. A ocorrência de duas variedades diferentes de DNA dentro do citoplasma de uma única célula é chamada de **heteroplasmia**. Quando uma célula heteroplásmica se divide, as organelas se separam de forma aleatória em duas células descendentes, em um processo chamado de **segregação replicativa** (**Figura 11.11**), e o acaso determina a proporção de organelas mutantes em cada célula. Embora a maioria das células descendentes herde uma mistura de organelas mutantes e normais, apenas o acaso pode determinar se algumas células recebem

organelas com apenas sequências mutantes ou apenas do tipo selvagem; essa situação, na qual todas as organelas são geneticamente idênticas, é conhecida como **homoplasmia**. É frequente ocorrer a fusão de mitocôndrias.



Conclusão: A maioria das células produzidas é heteroplásmica, mas, por acaso, algumas células recebem apenas um tipo de organela (p. ex., elas podem receber todas normais ou todas mutantes).

Figura 11.11 Organelas em uma célula heteroplásmica se dividem aleatoriamente nas células descendentes. Este diagrama ilustra a segregação replicativa na mitose; o mesmo processo também ocorre na meiose.

Quando a segregação replicativa ocorre nas células somáticas, ela pode criar variação fenotípica dentro de um único organismo; diferentes células do organismo podem ter diferentes proporções de sequências mutantes e do tipo selvagem, resultando em diversos graus de expressão fenotípica em tecidos variados. Quando essa segregação ocorre nas células germinativas de um doador com citoplasma heteroplásmico, podem surgir diferentes fenótipos entre os descendentes.

A doença conhecida como epilepsia mioclônica com fibras rotas vermelhas (MERRF) é causada por uma mutação em um gene mtDNA. Uma pessoa de 20 anos portadora dessa mutação em 85% de seus mtDNAs apresentava fenótipo normal, enquanto um primo que tinha a mutação em 96% de seus mtDNAs apresentava a forma grave da doença. A gravidade das doenças causadas por mutações no mtDNA está frequentemente relacionada com a proporção de sequências de mtDNA mutantes herdadas ao nascer. **Resolva o Problema 28**

Traços codificados pelo mtDNA. Foram estudados vários traços afetados pelo DNA da organela. Um dos primeiros a ser examinado em detalhes foi o fenótipo produzido por pequenas (*petite*) mutações na levedura. No final dos anos 1940, Boris Ephrussi e seus colegas observaram que algumas colônias de leveduras eram muito menores que o normal quando cresciam em meio sólido. O exame dessas colônias *petite* revelou que as taxas de crescimento das células nas colônias estavam muito reduzidas. Os resultados de estudos bioquímicos demonstraram que mutantes *petite* eram incapazes de fazer a respiração aeróbica; elas obtinham toda sua energia a partir do metabolismo anaeróbico (glicólise e fermentação), que é muito menos eficiente do que a respiração aeróbica e gera colônias menores.

Algumas mutações *petite* são defeitos no DNA nuclear, mas a maioria dessas mutações ocorre no DNA mitocondrial. Mutações *petite* mitocondriais têm grandes deleções no mtDNA, ou em alguns casos, o mtDNA está totalmente ausente. Parte do mtDNA codifica enzimas que catalisam a respiração aeróbica; portanto, as mutantes *petite* são incapazes de fazer a respiração aeróbica e não produzem quantidades normais de ATP, o que inibe seu crescimento.

Outro tipo de mutação no mtDNA conhecida ocorre na *Neurospora* (ver *Capítulo 15*). Isoladas por Mary Mitchell em 1952, as mutações *poky* crescem lentamente, apresentam herança citoplasmática, e têm quantidade anormal de citocromos. Os citocromos são componentes proteicos da cadeia de transporte de elétrons das mitocôndrias e exercem papel integral na produção de ATP. A maioria dos organismos tem três tipos primários de citocromos: citocromo *a*, citocromo *b* e citocromo *c*. As mutações *poky* têm o cromossomo *c*, mas não têm os cromossomos *a* ou *b*. Como as mutações *petite*, as mutações *poky* são defeituosas na síntese de ATP, e, portanto, as células crescem mais lentamente do que as do tipo selvagem. **Resolva o Problema 32**

Recentemente, foram identificadas várias doenças genéticas resultantes de mutações no mtDNA nos humanos. Além da síndrome MERRF mencionada anteriormente, a neuropatia óptica hereditária de Leber (LHON) resulta de mutações nos genes do mtDNA que codificam as proteínas envolvidas no transporte de elétrons. A LHON leva à perda súbita de visão na meia-idade. Outra doença causada por mutações no mtDNA é a fraqueza muscular neurogênica, ataxia e retinite pigmentosa (NARP), que é caracterizada por convulsões, demência e atraso no desenvolvimento. Outras doenças mitocondriais incluem a síndrome de Kearns-Sayre (KSS) e a oftalmoplegia externa crônica (CEOP); ambas levam à paralisia dos músculos oculares, pálpebras caídas e, em casos graves, perda de visão, surdez e demência. Todas essas doenças apresentam herança citoplasmática e expressão variável (ver *Capítulo 5*).

A esterilidade masculina citoplasmática, um fenótipo mutante encontrado em mais de 140 espécies diferentes de plantas e herdado apenas da genitora, é um traço nas plantas produzido por mutações nos genes mitocondriais. Essas mutações inibem o desenvolvimento do pólen, mas não afetam a fertilidade feminina.

Também foram descobertas várias mutações no cpDNA. Uma das primeiras a ser identificada foi a variação da folha na maravilha (*Mirabilis jalapa*), estudada por Carl Correns em 1909 (ver *Capítulo 5*). Na alga verde *Chlamydomonas*, as mutações resistentes à estreptomicina ocorrem no cpDNA, e vários mutantes que exibem pigmentação e crescimento alterados nas plantas superiores foram rastreados para encontrar defeitos no cpDNA.

Conceitos

Na maioria dos organismos, os genes codificados pelo mtDNA e cpDNA são herdados de um único genitor. Um gameta pode ter mais de um tipo diferente de mtDNA ou cpDNA; nesses casos, a segregação aleatória do DNA da organela produziria variação fenotípica em um único organismo ou pode produzir diferentes níveis de expressão fenotípica entre os descendentes.

Checagem dos conceitos 9

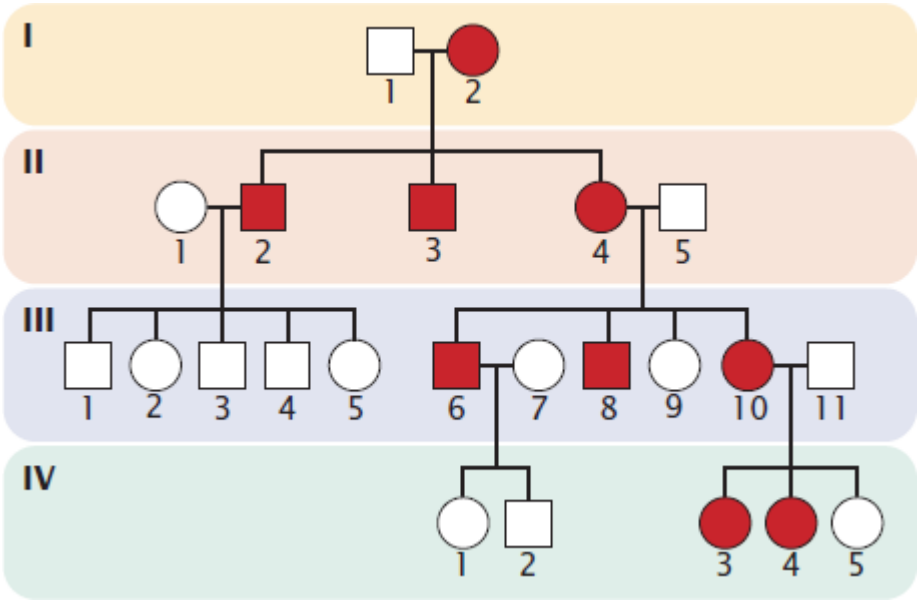
Em alguns organismos, os traços codificados pelo mtDNA podem ser herdados de ambos os genitores. Essa

observação indica que nesses organismos:

a. As mitocôndrias não exibem segregação replicativa.
b. Há heteroplasmia.
c. Espermatozoides e óvulos fornecem citoplasma para o zigoto.
d. Existem múltiplas cópias de mtDNA em cada célula.

Problema desenvolvido

Para ilustrar a herança de um traço codificado pelo DNA de organela, analise o problema a seguir. Um médico examina um homem jovem que tem um distúrbio muscular progressivo e anomalias visuais. Vários parentes do paciente têm a mesma doença, como apresentado no heredograma adiante. O grau de expressão do traço varia muito nos membros da família: alguns são muito pouco afetados, enquanto outros desenvolveram sintomas graves ainda jovens. O médico conclui que essa doença é causada por uma mutação no genoma mitocondrial. Você concorda com a conclusão do médico? Por que sim ou por que não? O distúrbio poderia ser causado por uma mutação em um gene que se encontra no núcleo da célula? Explique sua justificativa.



Estratégias para a solução

De quais informações você precisa para solucionar o problema?

Esse distúrbio poderia ser causado por uma mutação no genoma mitocondrial ou uma mutação em um gene nuclear e o motivo para tal.

Que informações são fornecidas para solucionar o problema?

- O homem jovem tem um distúrbio muscular progressivo e anomalias visuais
- Um heredograma ilustrando a família desse homem
- O traço varia muito entre os membros da família.

Etapas para a solução

A conclusão de que o distúrbio é causado por uma mutação no genoma mitocondrial é confirmada pelo heredograma e pela observação da expressão variável nos membros afetados da mesma família. O distúrbio é transmitido apenas pelas mães afetadas a seus filhos e filhas; quando os pais são afetados, nenhum dos filhos tem o traço (como observado nas crianças de II-2 e III-6). Esse desfecho é esperado quando os traços são determinados por mutações no mtDNA, pois as mitocôndrias estão no citoplasma e são herdadas apenas de um genitor (nos seres humanos, a mãe). O traço não pode ser recessivo ligado ao X, porque um cruzamento entre uma mulher com o traço (X^aX^a) e um homem sem o traço (X^+Y) *não* resultaria em filhas com o traço (X^aX^a), o que observamos em III-10, IV-3 e IV-4. Não pode ser dominante ligado ao X porque II-2 e III-6 teriam que transmitir para suas filhas, que não são afetadas (exceto se o traço exibido tiver penetrância incompleta).

O fato de alguns descendentes de mulheres afetadas não mostrarem o traço (III-9 e IV-5) e de a expressão variar de uma

pessoa para outra sugere que as pessoas afetadas são heteroplásmicas, com mitocôndrias mutantes e do tipo selvagem. A segregação aleatória das mitocôndrias na meiose pode produzir gametas com diferentes proporções de sequências mutantes e do tipo selvagem, resultando em diferentes graus da expressão fenotípica nos descendentes. Provavelmente, os sintomas do distúrbio surgem quando uma proporção mínima de mitocôndrias é mutante. Apenas por acaso, alguns dos gametas produzidos por uma mãe afetada têm poucas mitocôndrias mutantes e resultam em descendentes sem a doença.

Outra possível explicação para o distúrbio é que ele surge a partir de um gene autossômico dominante. Quando a pessoa afetada (heterozigota) acasala com uma pessoa não afetada (homozigota), é esperado que cerca da metade dos descendentes tenha o traço, mas, apenas por acaso, alguns pais afetados terão filhos não afetados pela doença. Os indivíduos II-2 e III-6 no heredograma poderiam ser homens e seu sexo poderia não estar relacionado com o modo de transmissão. A expressão variável poderia ser explicada pela expressividade variável (ver *Capítulo 5*).

Para ganhar mais experiência sobre a herança dos traços codificados pelas organelas, resolva o **Problema 29** no final do capítulo.

Genoma mitocondrial

Na maioria dos animais e fungos, o genoma mitocondrial inteiro existe em uma única molécula de DNA, circular e enrolada, embora possa haver muitas cópias desse genoma em cada célula. A estrutura circular da mitocôndria é semelhante à estrutura do cromossomo bacteriano. Os genomas mitocondriais das plantas existem como uma coleção complexa de múltiplas moléculas de DNA circulares. Em algumas espécies, o genoma mitocondrial é composto por uma única molécula de DNA linear.

Cada mitocôndria tem múltiplas cópias do genoma mitocondrial, e uma célula pode ter muitas mitocôndrias. Uma célula hepática de rato, por exemplo, tem de 5 a 10 moléculas de mtDNA em cada uma de suas 1.000 mitocôndrias; então, cada célula tem de 5.000 a 10.000 cópias do genoma mitocondrial. O DNA mitocondrial constitui cerca de 1% de todo DNA celular em uma célula hepática de rato. Como os cromossomos eubacterianos, o mtDNA não tem as proteínas histona normalmente associadas ao DNA nuclear eucariótico, embora seja complexado com outras proteínas com propriedades semelhantes às da histona. O teor de guanina-citosina (G-C) do mtDNA é muito diferente do teor no DNA nuclear, de modo que o mtDNA pode ser separado do DNA nuclear pela centrifugação por gradiente de densidade.

Os genomas mitocondriais são menores em comparação com os genomas nucleares, e com tamanho muito diferente entre os organismos (**Quadro 11.4**). O tamanho dos genomas mitocondriais da maioria das espécies varia de 15.000 pb a 65.000 pb; em algumas espécies eles são muito menores (p. ex., o genoma do *Plasmodium falciparum*, o parasita que causa malária, tem apenas 6.000 pb), enquanto os de algumas plantas são muitos milhões de pares de base. Embora a quantidade de DNA nos genomas mitocondriais varie muito, não existe correlação entre o tamanho do genoma e o número de genes. O número de genes é mais constante do que o tamanho do genoma; a maioria das espécies tem apenas 40 a 50 genes. Esses genes codificam cinco funções básicas: respiração e fosforilação oxidativa, tradução, transcrição, processamento do DNA e importação de proteínas para a célula. Grande parte da variação no tamanho dos genomas mitocondriais se deve a diferenças em sequências de DNA não codificadas. Como já foi mencionado, os genes para a maioria das proteínas e enzimas encontradas nas mitocôndrias são, de fato, codificados pelo DNA nuclear.

Quadro 11.4 Tamanhos de genomas mitocondriais em organismos selecionados.

Organismo	Tamanho do mtDNA (pb)
Pichia canadensis (fungo)	27.694
Podospora anserina (fungo)	100.314
Saccharomyces cerevisiae (fungo)	85.779*
Drosophila melanogaster (mosca-da-fruta)	19.517
Lumbricus terrestris (minhoca)	14.998
Xenopus laevis (sapo)	17.553
Mus musculus (camundongo doméstico)	16.295
Homo sapiens (humano)	16.569

Chlamydomonas reinhardtii (alga verde)	15.758
Plasmodium falciparum (protista)	5.966
Paramecium aurelia (protista)	40.469
Arabidopsis thaliana (planta)	166.924
Cucumis melo (planta)	2.400.000

*O tamanho varia entre as cepas.

mtDNA humano. O mtDNA humano é uma molécula circular com 16.569 pb que codifica dois rRNAs, 22 tRNAs e 13 proteínas. As duas cadeias de nucleotídios da molécula apresentam diferente composição de base: a cadeia pesada (H) tem mais nucleotídios guanina, enquanto a cadeia leve (L) tem mais nucleotídios citosina. A cadeia H é o molde para ambos os rRNAs, 14 dos 22 tRNAs e 12 das 13 proteínas, enquanto a cadeia L serve como molde para apenas 8 dos tRNAs e 1 proteína. A **alça D** (**Figura 11.12**) é uma região do mtDNA que tem sítios onde são iniciadas a replicação e a transcrição do mtDNA. O mtDNA humano tem uma organização muito econômica: há poucos nucleotídios não codificados entre os genes e quase todo o RNA mensageiro codifica proteínas. O mtDNA humano também tem muito pouco DNA repetitivo. A região do mtDNA humano que contém nucleotídios não codificadores é a alça D.

mtDNA de levedura. A organização do mtDNA de levedura é bem diferente da organização do mtDNA humano. Embora o genoma mitocondrial da levedura com 78.000 pb seja quase cinco vezes maior, ele codifica apenas seis genes adicionais, para um total de 2 rRNAs, 25 tRNAs e 16 polipeptídios (**Figura 11.13**). A maior parte do DNA extra no genoma mitocondrial da levedura tem sequências não codificadoras encontradas dentro e entre os genes.

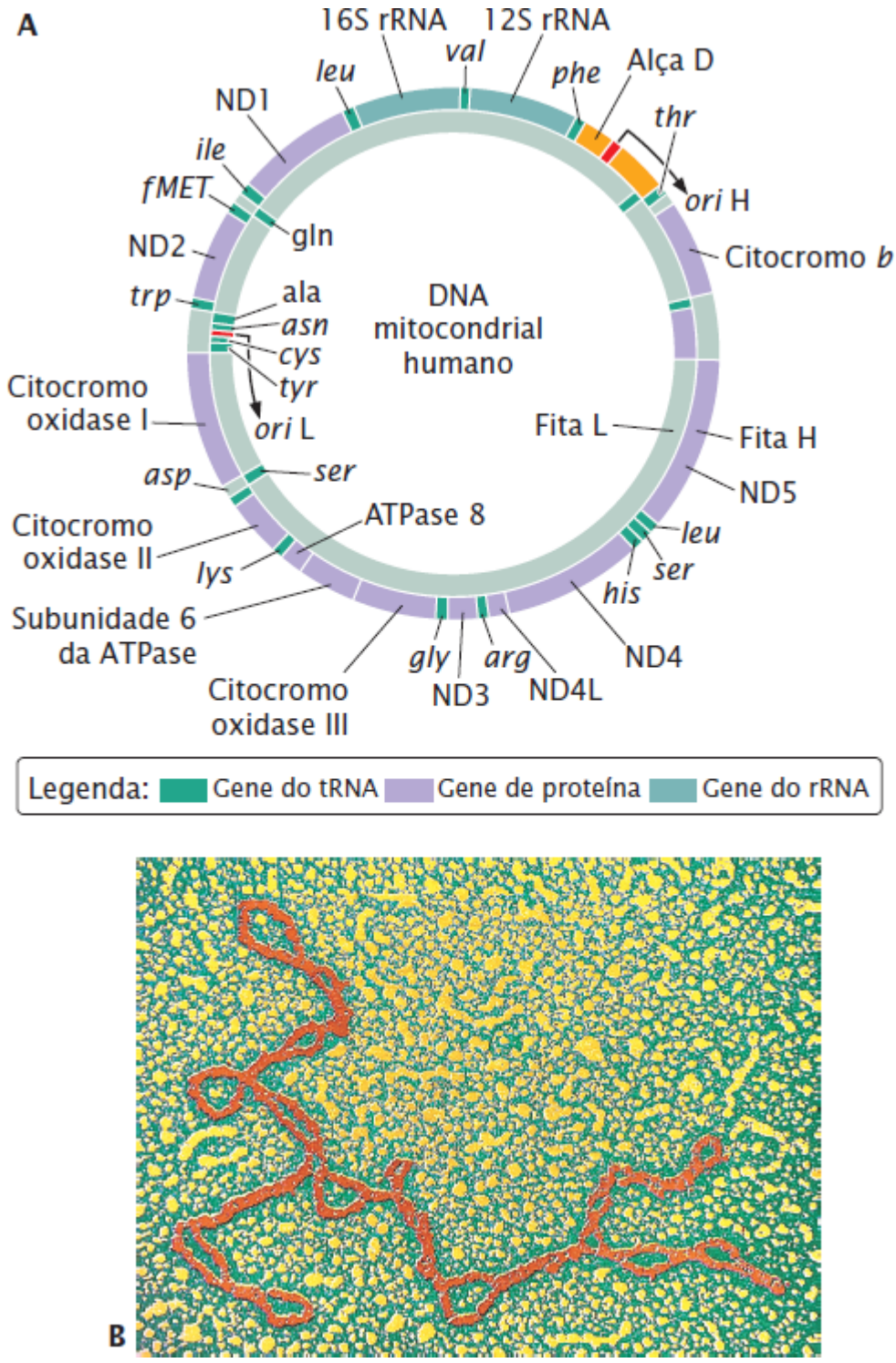


Figura 11.12 **O genoma mitocondrial humano tem 16.569 pb e apresenta uma organização muito econômica. A.** O círculo externo representa a cadeia pesada (H), e o círculo interno representa a cadeia leve (L). As origens de replicação para as cadeias H e L são *ori* H e *ori* L, respectivamente. *ND* identifica genes que codificam subunidades da NADH desidrogenase. **B.** Micrografia eletrônica de mtDNA isolado. (*Parte B*: CNRI/Photo Researchers.)

mtDNA de planta com inflorescência. As plantas florescentes (angiospermas) têm os maiores e mais complexos genomas mitocondriais conhecidos, que variam de 186.000 pb na mostarda branca a 2.400.000 pb no melão. Até espécies de plantas muito próximas podem apresentar grande diferença nos tamanhos de seus mtDNA.

Parte dessa enorme variação de tamanho no mtDNA das plantas florescentes pode ser explicada pela existência de longas sequências que são repetições diretas. O *crossing over* entre essas repetições pode gerar múltiplos cromossomos circulares de diferentes tamanhos. O genoma mitocondrial no nabo, por exemplo, é um "círculo principal" com 218.000 pb com repetições diretas. A recombinação homóloga entre as repetições pode gerar dois círculos menores de 135.000 pb e 83.000 pb (**Figura 11.14**). Outras espécies têm várias repetições diretas, fornecendo possibilidades para eventos complexos de *crossing over* que podem aumentar ou reduzir o número e tamanho dos círculos.

Conceitos

O genoma mitocondrial consiste em DNA circular sem proteínas histona associadas, embora seja complexado com outras proteínas que apresentam propriedades semelhantes às da histona. O tamanho e as estruturas do mtDNA são muito diferentes entre os organismos. O mtDNA humano exibe extrema economia, mas os mtDNAs encontrados na levedura e nas plantas florescentes têm muitos nucleotídios não codificadores e sequências repetitivas. O DNA mitocondrial na maioria das plantas florescentes é grande e tem uma ou várias grandes repetições diretas que podem recombinar-se para gerar moléculas menores ou maiores.

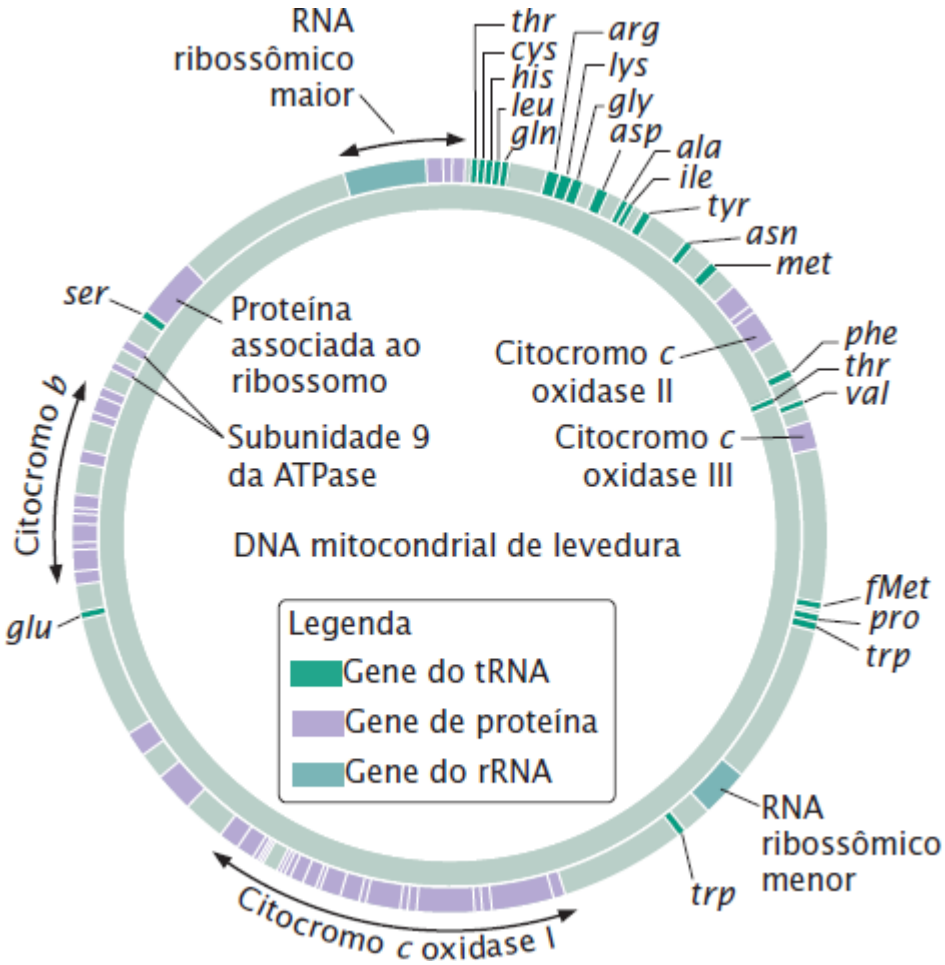


Figura 11.13 **O genoma mitocondrial da levedura, com 78.000 pb, tem muito DNA não codificador.**

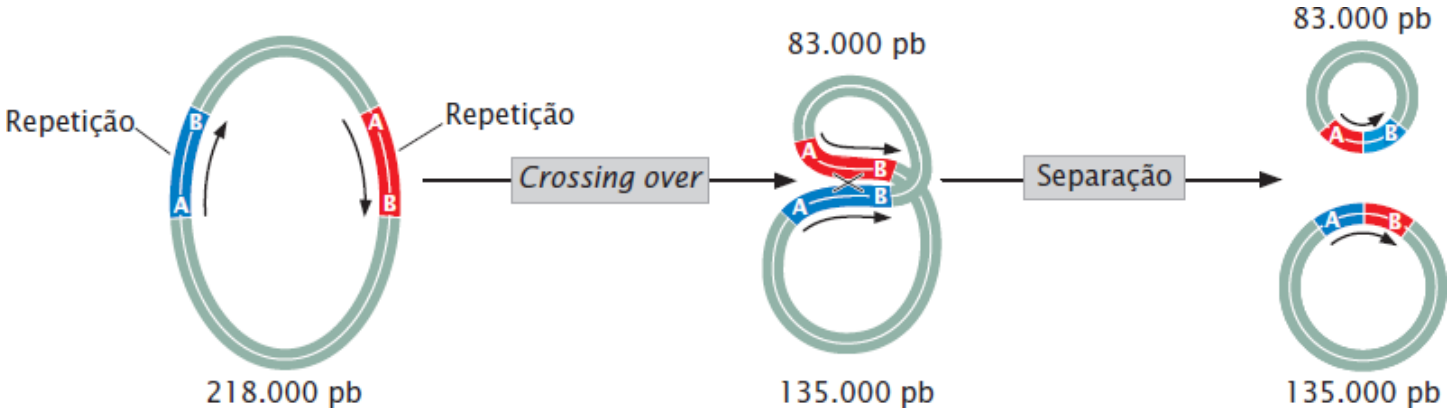


Figura 11.14 **A variação do tamanho no mtDNA de planta pode ser gerada pela recombinação entre repetições diretas**. No nabo, o genoma mitocondrial é composto por um "ciclo principal" de 218.000 pb; o *crossing over* entre as repetições diretas produz dois círculos menores de 135.000 pb e 83.000 pares de nucleotídios.

Evolução do DNA mitocondrial

Como já foi mencionado, as comparações das sequências de DNA mitocondrial com as sequências de DNA nas bactérias reforçam uma origem eubacteriana comum para todo mtDNA. Contudo, os padrões de evolução observados no mtDNA variam

muito entre os diferentes grupos de organismos.

As sequências de mtDNA vertebrado exibem uma taxa acelerada de evolução: as sequências no mtDNA de mamífero, por exemplo, em geral mudam de 5 a 10 vezes mais rápido do que o DNA nuclear de mamífero. A taxa acelerada de evolução observada no mtDNA de vertebrados é causada por sua elevada taxa de mutação, o que torna possível que as sequências de DNA mudem mais rapidamente. Apesar da alta taxa de evolução da sequência, o número de genes e a organização dos genomas mitocondriais nos vertebrados são relativamente constantes. Ao contrário, as sequências de mtDNA de plantas evoluem lentamente, em uma taxa de apenas 1/10 do que o genoma nuclear, mas seu conteúdo genético e sua organização mudam rapidamente. A razão para essas diferenças básicas nas taxas de evolução ainda é desconhecida.

O DNA mitocondrial foi muito estudado para reconstruir os padrões de evolução nos humanos e muitos outros organismos. Algumas das vantagens do mtDNA para estudar a evolução são: (1) o tamanho pequeno e a abundância de mtDNA na célula; (2) a rápida evolução das sequências de mtDNA em alguns organismos, o que facilita o estudo de grupos próximos e (3) a herança materna do mtDNA e a ausência de recombinação, o que possibilita rastrear linhas de descendência feminina. Foram analisadas amostras de mtDNA humano a partir de milhares de pessoas pertencentes a centenas de grupos étnicos diferentes pelo mundo. Essas amostras de mtDNA estão ajudando a desvendar muitos aspectos da evolução e história humana. Por exemplo, estudos iniciais de sequências de mtDNA levaram à proposta de que pequenos grupos de seres humanos migraram da África cerca de 85.000 anos atrás e povoaram o resto do mundo. Essa hipótese é chamada de *Out of Africa* ou *African Replacement*, e vem ganhando ampla aceitação. A hipótese *Out of Africa* está apoiada por estudos adicionais de sequências de DNA a partir do cromossomo Y e de genes nucleares. O uso de mtDNA em estudos sobre evolução será descrito com mais detalhes no Capítulo 26.

Na concepção, um zigoto de mamífero herda aproximadamente 100.000 cópias de mtDNA herdado do óvulo. Por causa do grande número de moléculas de mtDNA em cada célula e da alta taxa de mutação no mtDNA, seria esperado que a maioria das células tivesse uma mistura de moléculas de mtDNA mutantes e do tipo selvagem (heteroplasmia). Entretanto, é raro observar a heteroplasmia: na maioria dos organismos, as cópias do mtDNA são geneticamente idênticas (homoplasmia). Para explicar a uniformidade do mtDNA dentro dos organismos individuais, os geneticistas criaram uma hipótese de que, no início do desenvolvimento ou da formação do gameta, o mtDNA sofre algum tipo de gargalo durante o qual os mtDNAs de uma célula são reduzidos a algumas cópias, que então se replicam e dão origem a todas as cópias subsequentes de mtDNA. Por esse processo, a variação genética no mtDNA em uma célula é eliminada e a maioria das cópias é idêntica. Estudos recentes confirmam a existência do gargalo, mas há uma evidência contraditória sobre em que etapa do desenvolvimento ele surge.

Conceitos

Todo mtDNA parece ter evoluído a partir de um ancestral eubacteriano comum, mas os padrões de evolução observados em diferentes genomas mitocondriais variam muito. O mtDNA de vertebrados exibe mudança rápida na sequência, mas pouca mudança no conteúdo e na organização do gene, enquanto o mtDNA das plantas exibe pouca mudança na sequência, mas muita variação no conteúdo e na organização do gene. As sequências de DNA mitocondrial são usadas para estudar padrões de evolução.

Dano ao DNA mitocondrial e sua associação ao envelhecimento

Os sintomas de muitas doenças genéticas humanas causadas por defeitos no mtDNA surgem na meia-idade ou mais tarde, e sua gravidade aumenta com o avanço da idade. Uma hipótese para explicar esse fato está relacionada com o declínio na fosforilação oxidativa com o envelhecimento.

A fosforilação oxidativa é o processo que gera ATP, o principal transportador de energia na célula. Esse processo ocorre na membrana interna da mitocôndria e requer várias proteínas diferentes, algumas codificadas pelo mtDNA e outras pelos genes nucleares. A fosforilação oxidativa declina com a idade e, se cair abaixo de um limiar crítico, os tecidos não produzem ATP suficiente para manter as funções vitais e surgem sintomas de doenças. As pessoas iniciam a vida com uma capacidade em excesso para fosforilação oxidativa; essa capacidade reduz com a idade, mas a maioria das pessoas envelhece ou morre antes que o limiar crítico seja alcançado. Pessoas que nasceram com doenças mitocondriais carreiam mutações em seu mtDNA que reduzem a capacidade de fosforilação oxidativa. Ao nascer, sua capacidade pode ser suficiente para responder a suas necessidades de ATP, mas à medida que a capacidade de fosforilação oxidativa reduz com a idade, elas cruzam o limiar crítico e começam a apresentam sintomas da doença.

Por que a capacidade de fosforilação oxidativa reduz com a idade? Uma possível explicação é que o dano ao mtDNA se acumula com a idade: as deleções e substituições de base no mtDNA aumentam com a idade. Por exemplo, uma deleção comum de 5.000 pb no mtDNA não ocorre nas células do músculo cardíaco normal antes dos 40 anos, mas, posteriormente, essa deleção ocorre em uma frequência crescente. A mesma deleção é encontrada em uma frequência menor no tecido cerebral

normal antes dos 75 anos, mas ocorre em 11 a 12% dos mtDNAs nos núcleos da base aos 80 anos. Pessoas com doenças genéticas relacionadas com o mtDNA envelhecem prematuramente porque suas vidas já começaram com mtDNA danificado.

O mecanismo de aumento de danos ao DNA relacionados com a idade ainda é desconhecido. Os radicais de oxigênio, compostos altamente reativos que são subprodutos naturais da fosforilação oxidativa, danificam o DNA (ver *Capítulo 18*). Como o mtDNA está perto das enzimas envolvidas na fosforilação oxidativa, ele pode estar mais suscetível ao dano oxidativo do que o DNA nuclear. Quando o mtDNA é danificado, a capacidade de a célula produzir ATP cai.

Genoma de cloroplasto

Os geneticistas já identificaram que muitos traços associados aos cloroplastos exibem herança citoplasmática, indicando que esses traços não são codificados pelos genes nucleares. Em 1963, foi demonstrado que os cloroplastos tinham seu próprio DNA (**Figura 11.15**).

Entre diferentes plantas, o genoma de cloroplasto varia de 80.000 a 600.000 pares de base, mas a maior parte dos genomas de cloroplasto varia de 120.000 a 160.000 pb (**Quadro 11.5**). O DNA de cloroplasto é uma molécula única de fita dupla de DNA circular, altamente enrolada e sem as proteínas histona associadas. Como no mtDNA, são encontradas múltiplas cópias do genoma em cada cloroplasto e há múltiplas organelas por célula, então há muitas centenas a muitos milhares de cópias de cpDNA em uma típica célula de planta.

Quadro 11.5 Tamanhos de genomas de cloroplasto em organismos selecionados.

Organismo	Tamanho do cpDNA (pb)
Euglena gracilis (protista)	143.172
Porphyra purpurea (alga vermelha)	191.028
Chlorella vulgaris (alga verde)	150.613
Marchantia polymorpha (briófita)	121.024
Nicotiana tabacum (tabaco)	155.939
Zea mays (milho)	140.387
Pinus thunbergii (pinho-negro)	119.707

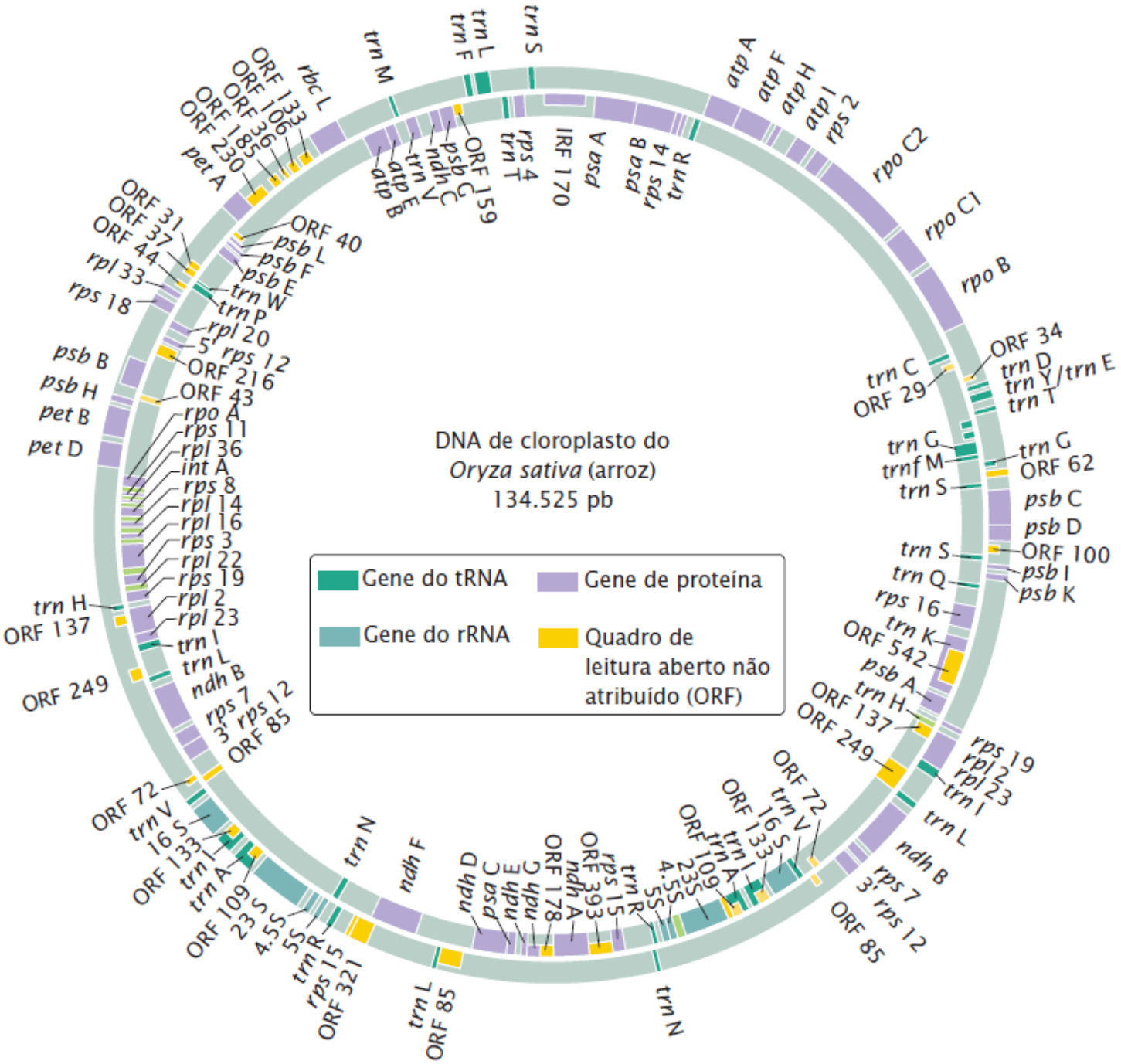


Figura 11.15 **DNA de cloroplasto do arroz.**

Foram sequenciados os genomas de cloroplasto oriundos de várias espécies de plantas e algas, e o cpDNA agora é identificado como sendo basicamente eubacteriano em sua organização: a ordem de alguns grupos de genes é a mesma que a observada na *E. coli* e muitos genes de cloroplasto estão organizados em grupos semelhantes aos encontrados nas bactérias. Várias das sequências no cpDNA são muito semelhantes às encontradas nos genes eubacterianos equivalentes.

Nas plantas vasculares, os cromossomos dos cloroplastos apresentam teor e ordem de genes semelhantes. Um típico genoma de cloroplasto codifica 4 genes de rRNA, de 30 a 35 genes de tRNA, várias proteínas de ribossomos, muitas proteínas envolvidas na fotossíntese e diversas proteínas com papéis em processos não fotossintéticos. Uma proteína importante codificada pelo cpDNA é a ribulose-1,5-bifosfato carboxilase-oxigenase (abreviada como RuBisCO), que participa na fixação do carbono na fotossíntese. A RuBisCO compõe cerca de 50% das proteínas encontradas nas plantas verdes e é considerada a proteína mais abundante na Terra. É uma proteína complexa composta por oito subunidades maiores idênticas e oito subunidades menores idênticas. A subunidade maior é codificada pelo DNA de cloroplasto, enquanto a subunidade menor é codificada pelo DNA nuclear. Grande parte do cpDNA é composta por sequências não codificadoras.

Evolução do DNA de cloroplasto. As sequências de DNA dos cloroplastos são muito semelhantes às encontradas nas cianobactérias (um grupo de bactérias fotossintéticas), então está claro que os genomas dos cloroplastos têm um ancestral eubacteriano. As sequências de cpDNA evoluíram lentamente em comparação com as sequências no DNA nuclear e algum mtDNA. Na maioria dos genomas de cloroplastos, o tamanho e a organização dos genes são semelhantes, embora existam

algumas exceções notáveis. Como eles evoluem mais lentamente e, como o mtDNA, são herdados de apenas um genitor, o cpDNA é útil para determinar as relações evolutivas entre diferentes espécies de plantas.

Conceitos

A maior parte dos genomas de cloroplasto é composta por uma única molécula de DNA circular, não complexada com proteínas histona. Embora exista uma considerável variação no tamanho entre as espécies, os cpDNAs encontrados na maioria das plantas vasculares têm cerca de 150.000 pb. As sequências de DNA de cloroplasto são semelhantes às sequências de DNA nas cianobactérias, o que apoia a teoria endossimbiótica.

Checagem dos conceitos 10

A organização do DNA de cloroplasto é muito semelhante à organização:

a. Das eubactérias.
b. De archaea.
c. Do DNA nuclear das plantas.
d. Do DNA nuclear dos eucariotos primitivos.

Transmissão da informação genética entre os genomas nucleares, mitocondriais e cloroplásticos ao longo da evolução

Muitas proteínas encontradas nas mitocôndrias e nos cloroplastos modernos são codificadas por genes nucleares, sugerindo que provavelmente parte do material genético original no endossimbionte tenha sido transferida para o núcleo. Essa hipótese é defendida pela observação de que algumas sequências de DNA normalmente encontradas no mtDNA foram detectadas no DNA nuclear de algumas cepas de levedura e milho. Do mesmo modo, sequências de cloroplasto foram encontradas no DNA nuclear do espinafre. Além disso, as sequências de genes nucleares que codificam as proteínas de organelas são muito semelhantes às sequências eubacterianas.

Também há indício de que o material genético foi deslocado dos cloroplastos para as mitocôndrias. Por exemplo, foram encontrados fragmentos de DNA para alguns genes de rRNA que são normalmente codificados no mtDNA do milho. Sequências a partir do gene que codifica a subunidade maior do RuBisCO, que é normalmente codificada pelo cpDNA, estão duplicadas no mtDNA do milho. Há, inclusive, evidências de que alguns genes nucleares foram deslocados para os genomas mitocondriais. A troca de material genético entre os genomas nucleares, mitocondriais e dos cloroplastos deu origem ao termo “DNA promíscuo” para descrever esse fenômeno. O mecanismo responsável por essa troca ainda não está totalmente elucidado.

Resumo dos conceitos

- Os cromossomos têm moléculas de DNA muito longas que estão firmemente comprimidas
- O superenrolamento resulta da tensão produzida quando rotações são adicionadas à molécula de DNA relaxada ou removidas desta. A rotação adicional produz o superenrolamento positivo; a falta de rotação produz o superenrolamento negativo. O superenrolamento é controlado por enzimas topoisomerases
- Um cromossomo bacteriano é composto por uma única molécula de DNA circular que é ligada a proteínas e existe como uma série de grandes alças. Em geral, aparece na célula como uma massa distinta conhecida como nucleoide
- Cada cromossomo eucariótico tem uma única molécula de DNA linear e longa, ligada a proteínas cromossômicas histona e não histona. A eucromatina sofre o ciclo normal de descondensação e condensação no ciclo celular. A heterocromatina permanece altamente condensada durante todo o ciclo celular
- O nucleossomo é composto por uma partícula central de oito proteínas histona e DNA que reveste o cerne. Os nucleossomos são dobrados em uma fibra de 30 nm que forma uma série de alças de 300 nm; essas alças estão ancoradas em suas bases pelas proteínas. As alças de 300 nm estão condensadas para formar uma fibra que por si só está altamente enrolada para produzir uma cromátide
- As regiões do cromossomo que estão sofrendo transcrição ativa são sensíveis à digestão pela DNase I, indicando que o DNA se abre durante a transcrição
- As alterações epigenéticas são modificações estáveis da expressão do gene que não precisam de mudanças nas sequências do DNA. Elas podem ocorrer por meio de alterações da estrutura da cromatina
- Os centrômeros são regiões cromossômicas onde as fibras do fuso se prendem; os cromossomos sem centrômeros são em geral

perdidos durante a divisão celular. A maioria dos centrômeros é definida por alterações epigenéticas na estrutura da cromatina. Os telômeros estabilizam as extremidades dos cromossomos

- O DNA eucariótico exibe três classes de sequências. Existem poucas cópias de DNA de sequência única. O DNA moderadamente repetitivo é composto por sequências moderadamente longas repetidas centenas a milhares de vezes. O DNA altamente repetitivo é composto por sequências muito curtas repetidas em *tandem* de muitos milhares a milhões de vezes
- As mitocôndrias e os cloroplastos são organelas eucarióticas com DNA próprio. A teoria endossimbiótica propõe que as mitocôndrias e os cloroplastos se originaram como organismos procarióticos de vida livre (especificamente eubactérias) que entraram em uma associação benéfica com as células eucarióticas
- Os traços codificados por mtDNA e cpDNA em geral são herdados a partir de um único genitor, em sua maioria, a mãe. A separação aleatória das organelas na divisão celular pode produzir variação fenotípica entre as células em um organismo e os descendentes de uma única fêmea
- O genoma mitocondrial em geral é composto por uma única molécula de DNA circular sem proteínas histona. O DNA mitocondrial apresenta diferença de tamanho nos diferentes grupos de organismos. O mtDNA humano é muito "econômico", com poucos nucleotídios não codificadores. Os mtDNAs de fungos e plantas têm muito DNA não codificador entre os genes
- As comparações de sequências de mtDNA sugerem que as mitocôndrias evoluíram de um ancestral eubacteriano. O mtDNA de vertebrados exibe mudança rápida na sequência, mas pouca mudança no conteúdo e na organização dos genes. O mtDNA de plantas exibe pouca mudança na sequência, mas muita variação no conteúdo e na organização dos genes
- As sequências de DNA mitocondrial são muito usadas para estudar a evolução
- Os genomas de cloroplasto são compostos por uma única molécula de DNA circular sem histonas e com pouca variação do tamanho. Cada célula de planta tem múltiplas cópias de cpDNA. As sequências de DNA de cloroplasto são as mais semelhantes às sequências encontradas nas cianobactérias e tendem a evoluir lentamente
- Por meio da evolução, muitos genes de mitocôndrias e cloroplastos se deslocaram para cromossomos nucleares. Em algumas plantas, há evidências de que as cópias dos genes de cloroplastos se deslocaram para o genoma mitocondrial.

Termos importantes

Alça D
Alterações epigenéticas
Cromossomo politênico
Desnaturação
DNA altamente repetitivo
DNA de cloroplasto (cpDNA)
DNA de sequência única
DNA ligador
DNA mitocondrial (mtDNA)
DNA moderadamente repetitivo
DNA repetitivo
Elemento curto intercalado (SINE)
Elemento longo intercalado (LINE)
Estado relaxado do DNA
Eucromatina
Família de genes
Heterocromatina
Heteroplasmia
Hibridização
Homoplasmia
Nucleoide
Nucleossomo
Paradoxo do valor de C
Proteína cromossômica não histona
Puff cromossômico
Renaturação (reanelamento)
Segregação replicativa

Sequência repetida intercalada
Sequência telomérica
Sequências de repetição em *tandem*
Shelterina
Superenrolamento
Superenrolamento negativo
Superenrolamento positivo
Temperatura de desnaturação (T_m)
Teoria endossimbiótica
Topoisomerase
Valor de C

Respostas da Checagem dos conceitos

1. b.
2. O DNA bacteriano não é complexado a proteínas histona e é circular.
3. b.
4. d.
5. Um cromossomo que perde seu centrômero não se separa no núcleo durante a mitose e em geral é perdido.
6. d.
7. a.
8. Muitos protistas modernos são hospedeiros de bactérias endossimbióticas. As mitocôndrias e os cloroplastos têm tamanho semelhante ao das eubactérias e têm seu próprio DNA, assim como os ribossomos têm tamanho e formato semelhantes aos ribossomos eubacterianos. Os antibióticos que inibem a síntese de proteínas nas eubactérias também inibem a síntese de proteínas nas mitocôndrias e nos cloroplastos. As sequências de genes do mtDNA e cpDNA são muito semelhantes às sequências de DNA eubacterianos.
9. c.
10. a.

Problemas desenvolvidos

Problema 1

Uma célula diploide de planta tem 2 bilhões de pares de base de DNA.

a. Quantos nucleossomos há na célula?
b. Indique o número de moléculas de cada tipo de proteína histona associada ao DNA genômico.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
O número de nucleossomos por célula e o número de cada tipo de proteína histona associada ao DNA.

Quais informações são fornecidas para solucionar o problema?
A célula tem 2 bilhões de pares de base de DNA.

Para a solução deste problema, revise:
Nucleossomo, na Seção 11.1.

Etapas para a solução

Cada nucleossomo abrange cerca de 200 pb de DNA: de 145 a 147 pb de DNA enrolado ao redor do cerne de histona, de 20 a 22 pb de DNA associado à proteína H1 e outros 30 a 40 pb de DNA ligador.

> Lembrete: A unidade repetida do cromossomo é um nucleossomo, que é composta por DNA complexado a proteínas histona.

a. Para determinar quantos nucleossomos há na célula, dividimos o número total de pares de base de DNA (2

× 10^9 pb) pelo número de pares de base por nucleossomo:

$$\frac{2 \times 10^9 \text{ nucleotídios}}{2 \times 10^2 \text{ nucleotídios por nucleossomo}} = 1 \times 10^7 \text{ nucleossomos}$$

Assim, existem aproximadamente 10 milhões de nucleossomos na célula.

b. Cada nucleossomo inclui duas moléculas de histonas H2A, H2B, H3 e H4. Portanto, existem 2 × 10^7 moléculas de histonas H2A, H2B, H3 e H4. Cada nucleossomo tem uma cópia de histona H1 associada; então, há 1 × 10^7 moléculas de H1.

Problema 2

Imagine que foi descoberta uma nova organela em um grupo obscuro de protistas. Essa organela tem um pequeno genoma de DNA, e alguns cientistas estão argumentando que, como os cloroplastos e mitocôndrias, essa organela se originou de uma eubactéria livre que entrou em uma relação endossimbiótica com o protista. Esquematize um plano de pesquisa para determinar se a nova organela evoluiu a partir de uma eubactéria livre. Que tipos de dados você coletaria e quais previsões você faria se a teoria estivesse correta?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Um plano de pesquisa com os tipos de dados que você coletaria e previsões.

Quais informações são fornecidas para solucionar o problema?

- Foi descoberta uma nova organela
- A organela tem um pequeno genoma de DNA
- A organela pode ter evoluído a partir de uma relação endossimbiótica.

Para a solução deste problema, revise:
Teoria endossimbiótica, *Genoma mitocondrial* e *Genoma de cloroplasto*, na Seção 11.4.

Etapas para a solução

Devemos examinar a estrutura, a organização e as sequências do genoma da organela. Se a organela mostrar apenas características de DNA eucariótico, então é provável que sua origem seja eucariótica, mas se ela apresentar algumas características de DNA eubacteriano, então esse fato apoia a teoria de origem bacteriana.

Poderíamos começar examinando as características gerais do DNA da organela. Se ele tiver uma origem eubacteriana, podemos esperar que fosse composto por uma molécula circular sem proteínas histona. Podemos comparar as sequências de DNA encontradas no seu genoma com as sequências homólogas dos genomas eubacterianos e eucarióticos. Se a teoria de uma origem endossimbiótica estiver correta, então as sequências da organela devem ser mais semelhantes às sequências homólogas encontradas nas eubactérias.

Lembrete: A teoria endossimbiótica propõe que as organelas evoluíram a partir das eubactérias.

Questões de compreensão

Seção 11.1

1. Como surge o superenrolamento? Qual é a diferença entre os superenrolamentos positivo e negativo?
2. Quais são as funções do superenrolamento na célula?
3. Descreva a composição e a estrutura do nucleossomo.
4. Descreva em etapas como a dupla-hélice do DNA, que tem 2 nm, dá origem a um cromossomo com 700 nm.

5. O que são cromossomos politênicos e *puffs* cromossômicos?
6. O que são alterações epigenéticas e como elas surgem?

Seção 11.2

7. Descreva a função do centrômero. Qual é a diferença do centrômero em relação às outras regiões do cromossomo?
8. Descreva a função e a estrutura molecular de um telômero.
9. Qual é a diferença entre eucromatina e heterocromatina?

Seção 11.3

10. O que é o valor de C de um organismo?
11. Descreva as diferentes classes de variação de sequência de DNA que existem nos eucariotos.

Seção 11.4

12. Explique por que muitos traços codificados por mtDNA e cpDNA exibem considerável variação em sua expressão, mesmo em membros da mesma família.
13. O que é teoria endossimbiótica? Como ela ajuda a explicar algumas das características das mitocôndrias e dos cloroplastos?
14. Que evidência apoia a teoria endossimbiótica?
15. Descreva brevemente a organização dos genes no genoma de cloroplasto.
16. O que significa o termo “DNA promíscuo”?

Questões e problemas aplicados

Introdução

17. Na introdução deste capítulo foi apresentado um estudo do comprimento do telômero nas crianças romenas. O estudo demonstrou que as crianças que cresceram em orfanatos tinham telômeros mais curtos do que aquelas que foram adotadas. Que efeito, se há algum, você acha que os telômeros mais curtos na infância poderiam ter na vida adulta?

Seção 11.1

*18. Compare e diferencie os cromossomos de procariotos daqueles dos eucariotos. Quais são suas semelhanças e diferenças?
*19. A. Você esperaria encontrar mais moléculas de histona H1 ou mais moléculas de histona H2 em uma típica célula eucariótica? Explique sua resposta. B. Você esperaria encontrar mais moléculas de H2A ou mais moléculas de H3? Explique sua resposta.
*20. Com base na sensibilidade do DNA à DNase I, ilustrada na Figura 11.6, que tipo de hemoglobina de galinha (embrionária ou adulta) é provavelmente produzida em maior quantidade nos tecidos e estados de desenvolvimento a seguir?
    a. Eritroblastos durante as primeiras 24 h.
    b. Eritroblastos no dia 5.
    c. Eritroblastos no dia 14.
    d. Células do cérebro durante o desenvolvimento.
21. Suponha que um pesquisador adicionou rapidamente uridina marcada com radioatividade na larva de *Drosophila*, cujo cromossomo politênico é apresentado na Figura 11.5. Marque na figura onde você esperaria ver o acúmulo de uridina radioativa.
*22. Uma célula diploide humana tem 6,4 bilhões de pares de base de DNA.
    a. Quantos nucleossomos há nessa célula? (Considere que o DNA ligador abrange 40 pb.)
    b. Quantas histonas estão complexadas a esse DNA?

*23. Você esperaria ver mais ou menos acetilação nas regiões do DNA que são sensíveis à digestão pela DNase I? Por quê?
24. Gunter Korge examinou várias proteínas que são secretadas pelas glândulas salivares da *Drosophila melanogaster* durante o desenvolvimento da larva (G. Korge. 1975. *Proceedings of the National Academy of Sciences of the United States of America* 72:4550-4554). Uma proteína, chamada fração 4 de proteína, era codificada por um gene encontrado no mapeamento de deleção como localizado no cromossomo X na posição 3C. Korge observou que, cerca de 5 h após a primeira síntese da fração 4 de proteína, uma região expandida e inchada formou-se no cromossomo X na posição 3C. Esse *puff* cromossômico desapareceu antes do final do estágio de larva, quando cessou a síntese da fração 4 da proteína. Ele observou que não havia *puff* na posição 3C em uma cepa especial de moscas que não secretavam a fração 4 de

proteína. Explique esses resultados. O que é o *puff* cromossômico na região 3 e por que seu aparecimento e desaparecimento coincidem abruptamente com a secreção da fração 4 da proteína?

25. Suponha que um químico desenvolveu um novo fármaco que neutraliza as cargas elétricas positivas das caudas das histonas. Qual seria o efeito mais provável deste novo fármaco na estrutura da cromatina? Este fármaco teria algum efeito na expressão do gene? Explique suas respostas.

Seção 11.3

*26. Qual das duas moléculas de DNA a seguir tem a menor temperatura de desnaturação menor? Por quê?

AGTTACTAAAGCAATACATC
TCAATGATTTCGTTATGTAG

AGGCGGGTAGGCACCCTTA
TCCGCCCATCCGTGGGAAT

ANÁLISE DE DADOS

27. Em um estudo de hibridização do DNA, o DNA foi isolado de uma espécie particular, marcada com ^{32}P e cortada em fragmentos pequenos (S. K. Dutta *et al*. 1967. *Genetics* 57:719-727). Foram comparadas as hibridizações entre esses fragmentos marcados e DNA desnaturado a partir de diferentes espécies. O quadro a seguir fornece as porcentagens de DNA marcado de trigo que hibridizou com moléculas de DNA de trigo, milho, rabanete e repolho.

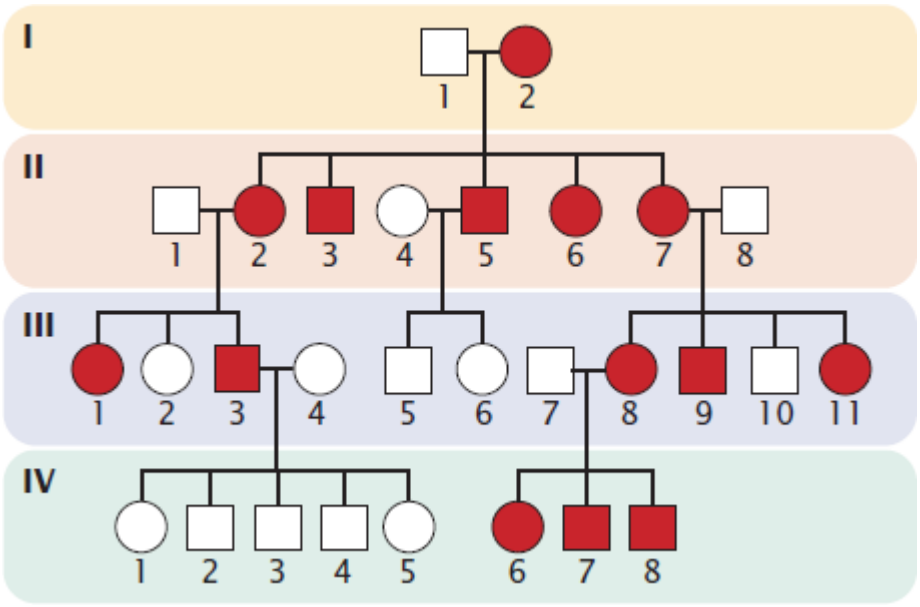
Espécie	**Porcentagem de DNA de trigo ligado hibridizado em relação ao trigo**
Trigo	100
Repolho	23
Milho	63
Rabanete	30

O que esses resultados indicam sobre as diferenças evolutivas entre esses organismos?

Seção 11.4

*28. Foi encontrada uma variação do trigo de cor verde-clara crescendo em um campo. A análise bioquímica revela que os cloroplastos nessa planta produzem apenas 50% da clorofila normalmente encontrada nos cloroplastos do trigo. Proponha um conjunto de cruzamentos para determinar se o fenótipo verde-claro é causado por uma mutação em um gene nuclear ou em um gene de cloroplasto.

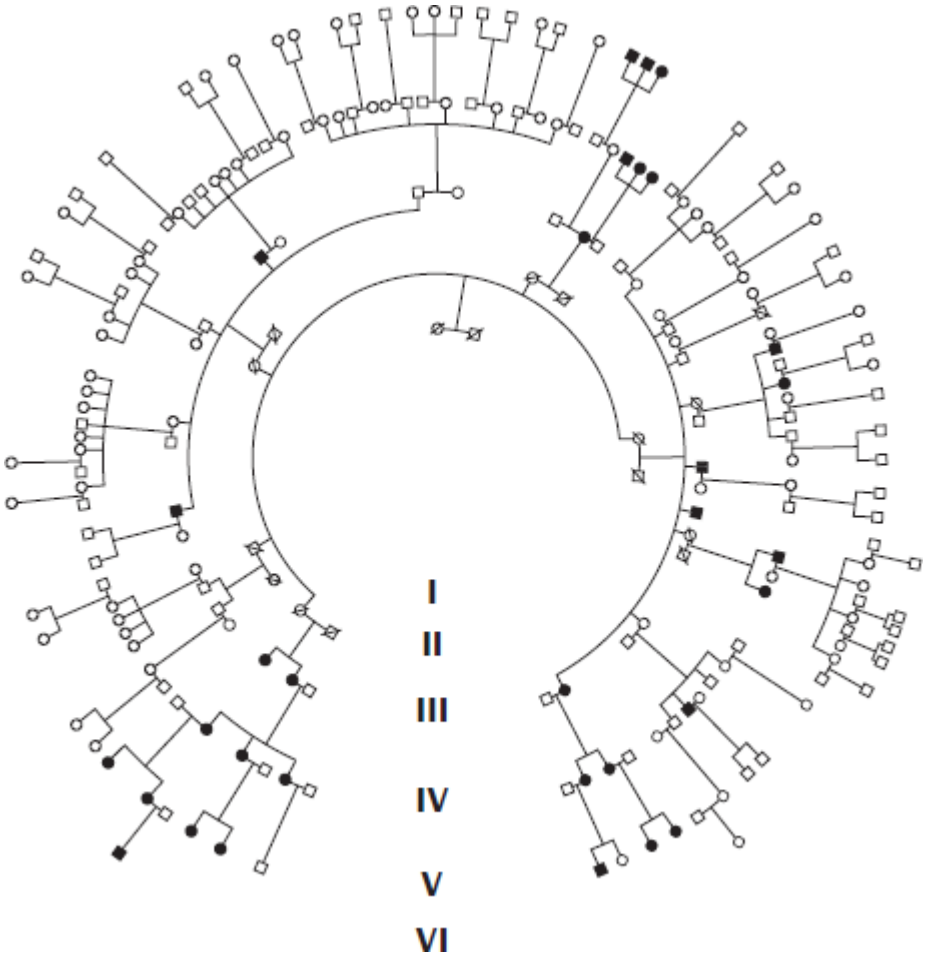
*29. Uma rara doença neurológica foi encontrada na família ilustrada no heredograma a seguir. Qual é a forma mais provável de herança para essa doença? Explique sua justificativa.



30. Considere que o distúrbio apresentado no heredograma do Problema desenvolvido nas pp. 280-281 é uma rara doença que surge a partir de um defeito no DNA mitocondrial. Se o indivíduo III-8 tiver uma filha, qual é a probabilidade de a filha herdar o distúrbio muscular do genitor acometido?

ANÁLISE DE DADOS

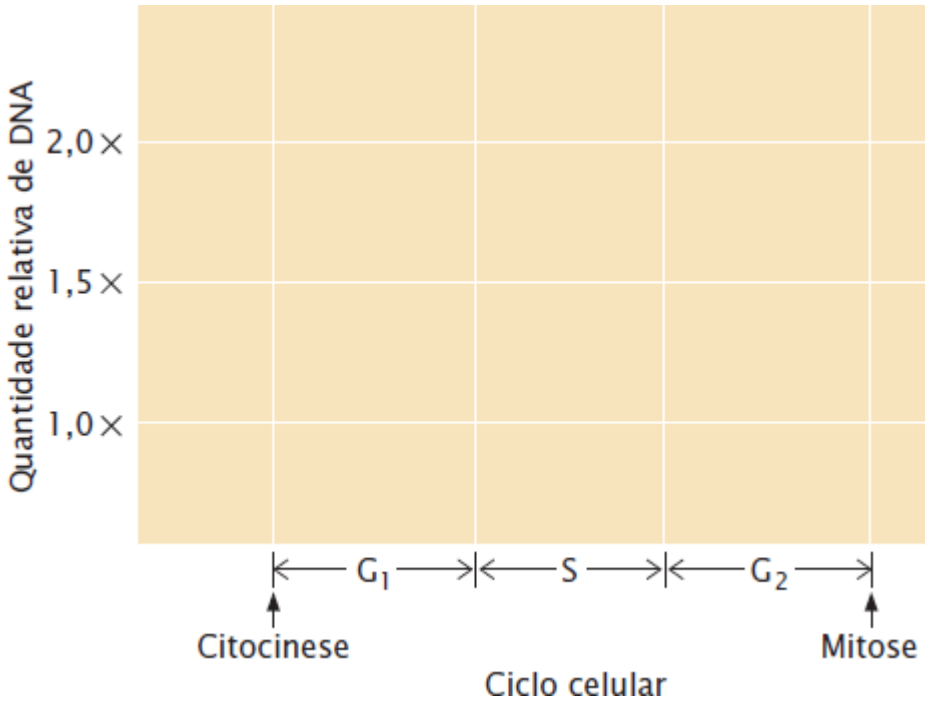
31. Fredrick Wilson e seus colegas estudaram membros de uma família grande que tinha baixos níveis sanguíneos de magnésio (ver heredograma a seguir). Eles argumentam que essa alteração de magnésio (associada a pressão sanguínea e colesterol elevados) é causada por uma mutação no mtDNA (F. H. Wilson *et al*. 2004. *Science* 306:1190-1194).
    a. Que evidência sugere que um gene no mtDNA é o responsável pelo distúrbio?
    b. Esse distúrbio poderia ser causado por um gene autossômico dominante? Por que sim ou por que não?



(De F. H. Wilson et al. 2004. Science 306:1190-1194.)

*32. Em determinada cepa de *Neurospora*, uma mutação *poky* exibe herança biparental, enquanto as mutações *poky* em outras cepas são herdadas apenas da genitora. Explique esses resultados.

*33. Uma cientista coleta células em vários pontos no ciclo celular e isola o DNA delas. Usando a centrifugação por gradiente de densidade, ela separa o DNA nuclear e o mtDNA. Então, mede a quantidade de mtDNA e de DNA nuclear em diferentes pontos no ciclo celular. No gráfico a seguir, desenhe uma linha para representar as quantidades relativas de DNA nuclear que você espera que a cientista encontre em cada célula durante todo o ciclo celular. Depois, desenhe uma linha pontilhada no mesmo gráfico para indicar a quantidade relativa de mtDNA que você esperaria ver em diferentes pontos durante todo o ciclo celular.



34. Em 1979, foi demonstrado que os ossos encontrados fora de Ekaterinburg, Rússia, eram do Czar Nicholas e de sua família, que foram executados em 1918 por um pelotão de fuzilamento bolchevique na Revolução Russa (ver a introdução do *Capítulo 14*). Para provar que os esqueletos eram da família real, foi extraído mtDNA das amostras de osso, ampliado por PCR e comparado com o mtDNA dos parentes vivos da família do czar.
    a. Por que foi analisado o DNA das mitocôndrias em vez do DNA nuclear? Quais são algumas das vantagens em usar o mtDNA para esse tipo de estudo?
    b. O DNA mitocondrial dos parentes vivos forneceria informações úteis para verificar se os esqueletos são da família real?

35. Antibióticos como cloranfenicol, tetraciclina e eritromicina inibem a síntese de proteínas nas eubactérias, mas não têm efeito na síntese de proteínas codificadas pelos genes nucleares. A cicloeximida inibe a síntese de proteínas codificadas pelos genes nucleares, mas não tem efeito na síntese das proteínas eubacterianas. Como esses compostos poderiam ser usados para determinar quais proteínas são codificadas pelos genomas mitocondriais e de cloroplasto?

Questões desafiadoras

Seção 11.1

36. Um explorador descobre uma espécie nova e estranha de planta e envia parte do tecido da planta para uma geneticista estudar. A geneticista isola a cromatina da planta e a examina com um microscópio eletrônico. Ela observa o que parece ser um cordão de pérolas. Então, adiciona uma pequena quantidade de nuclease, que rompe o fio e libera as pérolas com 280 pb de DNA. Após a digestão com mais nuclease, um fragmento de 120 pb de DNA permanece preso ao cerne de proteínas histona. A análise do cerne de histonas revela as histonas nas seguintes proporções:

H1	12,5%
H2A	25%
H2B	25%
H3	0%
H4	25%

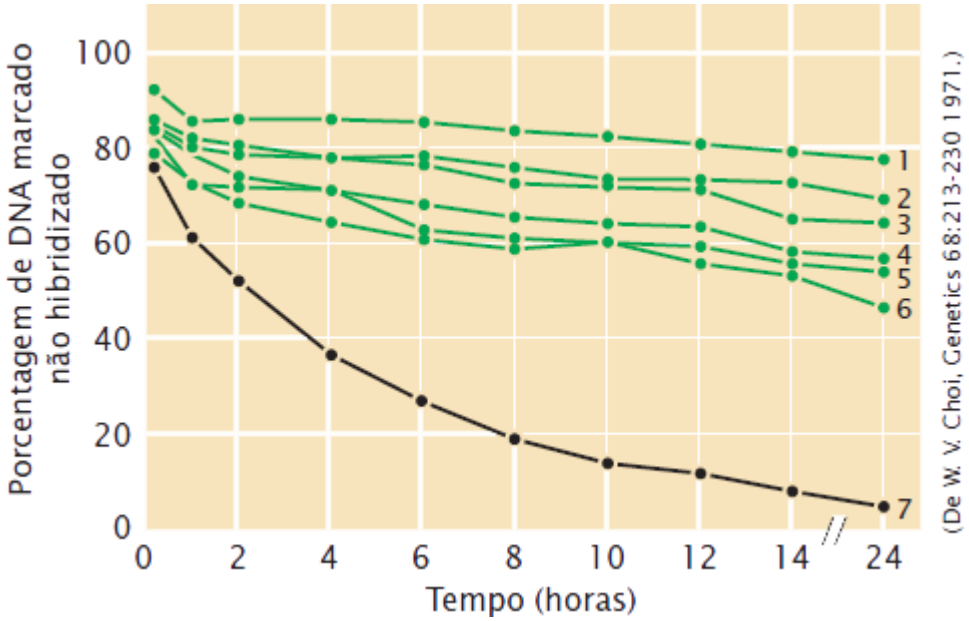
H7 (uma nova histona) 12,5%

Com base nessas observações, que conclusões a geneticista pode tirar sobre a provável estrutura do nucleossomo na cromatina dessa planta?

Seção 11.3

ANÁLISE DE DADOS

37. Nos experimentos de hibridização de DNA em seis espécies de plantas do gênero *Vicia*, o DNA foi isolado de cada uma das seis espécies, desnaturado com aquecimento e cortado em pequenos fragmentos (W. Y. Chooi. 1971. *Genetics* 68:213-230). Em um experimento, o DNA de cada espécie e da *E. coli* foi renaturado. O gráfico mostra os resultados desse experimento de renaturação.



Legenda
1 = V. melanops, 2 = V. sativa, 3 = V. benghalensis,
4 = V. atropurpurea, 5 = V. faba, 6 = V. narbonensis, 7 = E. coli

a. Você pode explicar por que o DNA da *E. coli* renatura muito mais rapidamente do que o DNA de todas as espécies de *Vicia*?

Ervilhaca (Vicia sativa). (Bob Gibbons/Alamy.)

b. Observe que, para as espécies *Vicia*, a velocidade de renaturação é muito maior na primeira hora e então reduz. O que poderia causar essa renaturação inicial rápida e a subsequente redução de velocidade?

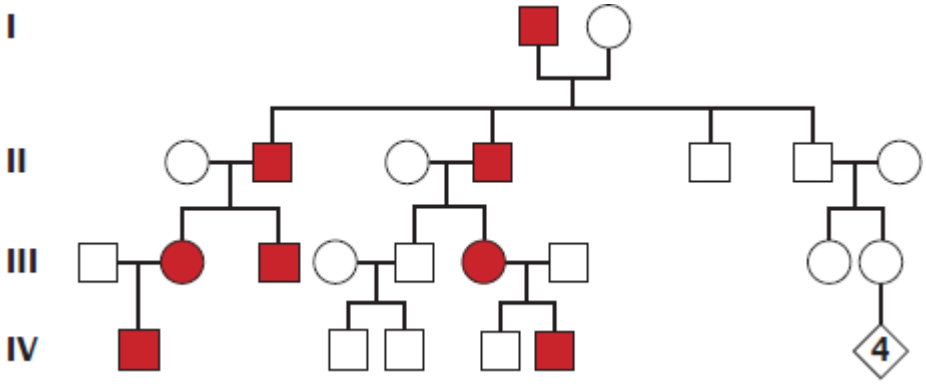
Seção 11.4

ANÁLISE DE DADOS

38. Steven Frank e Laurence Hurst argumentaram que uma mutação de herança citoplasmática nos seres humanos que tem efeitos graves nos homens, mas não afeta as mulheres, não será eliminada de uma população por meio da seleção natural, uma vez que apenas as mulheres transmitem o mtDNA (S. A. Frank and L. D. Hurst. 1996. *Nature* 383:224). Com base nesse argumento, explique por que os homens com a neuropatia óptica hereditária de Leber são afetados de forma mais grave do que as mulheres.

ANÁLISE DE DADOS

39. Em um estudo sobre uma miopatia, várias famílias exibiam problemas de visão, fraqueza muscular e surdez (M. Zeviani *et al*. 1990. *American Journal of Human Genetics* 47:904-914). A análise do mtDNA de pessoas afetadas nessas famílias revelou que uma grande quantidade de seu mtDNA tinha deleções de tamanhos variáveis. Diferentes membros da mesma família e até diferentes mitocôndrias da mesma pessoa apresentavam deleções de variados tamanhos, então o defeito envolvido parecia ser uma tendência de o mtDNA das pessoas afetadas ter as deleções. Um heredograma de uma das famílias estudadas é apresentado aqui. Os pesquisadores concluíram que esse distúrbio é herdado como um traço autossômico dominante, e então mapearam o gene responsável pela doença em uma posição no cromossomo 10 no núcleo.
    a. Quais características do heredograma descartam a herança de um traço codificado por um gene no mtDNA?
    b. Explique como uma mutação em um gene nuclear poderia levar a deleções no mtDNA.



(De M. Zeviani et al. 1990. American Journal of Human Genetics 47:904-914.)

40. Sequências de DNA mitocondrial foram detectadas em genomas nucleares de muitos organismos e, às vezes, são encontradas sequências de cpDNA no genoma mitocondrial. Proponha um mecanismo para como o “DNA promíscuo” poderia ter se deslocado entre os genomas nuclear, mitocondrial e cloroplástico.

12 Replicação e Recombinação de DNA

Topoisomerase, replicação e câncer

Em 1966, Monroe Wall e Mansukh Wani encontraram na casca da árvore-feliz (*Camptotheca acuminata*), uma planta rara nativa da China, uma cura potencial para o câncer. Wall e Wani trabalhavam na seleção de várias substâncias naturais com atividade anticâncer, na esperança de encontrar substâncias químicas que pudessem ter eficácia comprovada no tratamento da doença. Eles descobriram que um extrato dessa árvore era efetivo no tratamento de leucemia em camundongos. Por meio de análise química, eles conseguiram isolar o composto ativo, que foi batizado de camptotecina.

Na década de 1970, os médicos administravam camptotecina a pacientes com cânceres incuráveis. Embora o fármaco mostrasse atividade anticâncer, ele tinha efeitos colaterais tóxicos. Os químicos acabaram sintetizando vários análogos da camptotecina que eram menos tóxicos e mais eficientes no tratamento do câncer. Dois desses análogos, topotecano e irinotecano, são usados hoje em dia para o tratamento de cânceres de ovário, das células pequenas do pulmão e do cólon.

Por muitos anos, o mecanismo de inibição do câncer pelos derivados da camptotecina era desconhecido. Em 1985, quase 20 anos após sua descoberta, cientistas da Universidade de Johns Hopkins e dos laboratórios Smith Kline and French Laboratories (atualmente GlaxoSmithKline) mostraram que a camptotecina inibia um componente importante do maquinário da síntese do DNA nos seres humanos, uma enzima chamada topoisomerase I.

A quimioterapia para o câncer é uma tarefa delicada porque as células-alvo são do próprio paciente, e os fármacos precisam matar apenas as células do câncer, preservando a vida do paciente. Uma das características do câncer é a proliferação: a divisão das células do câncer é desregulada e muitas células se dividem muito rapidamente, dando origem a tumores capazes de crescer e se disseminar. Como aprendemos no Capítulo 2, antes que uma célula possa se dividir, ela precisa replicar seu DNA de modo que cada célula-filha receba uma cópia exata do material genético. Os pontos de verificação no ciclo celular garantem que a divisão celular não prossiga se a replicação do DNA for inibida ou defeituosa, e muitos tratamentos de câncer têm como foco a replicação do DNA.

A replicação do DNA é um processo complexo que requer vários componentes, e suas ações têm de estar intrinsecamente coordenadas a fim de garantir que o DNA seja copiado com fidelidade. Um componente essencial da replicação é a topoisomerase. À medida que o DNA se desenrola durante a replicação, tensão se acumula na frente da separação e as duas fitas se torcem ao redor uma da outra, como um nó de corda à medida que você afasta suas fitas. Essa torção do DNA é chamada superenrolamento (ver *Capítulo 11*). Os superenrolamentos, se não são removidos, acabam interrompendo a separação da fita, e a replicação é pausada.

As enzimas topoisomerase removem os superenrolamentos ao segurar firmemente o DNA e romper uma ou ambas as fitas. As fitas então rodam ao redor uma da outra, removendo o superenrolamento e a tensão. Após o DNA estar relaxado, a topoisomerase que removeu os superenrolamentos fecha novamente as extremidades rompidas do DNA.

A camptotecina interfere na topoisomerase I. O fármaco se insere no espaço criado pela ruptura da fita de DNA, impedindo a topoisomerase de fechar novamente as extremidades rompidas. Originalmente, os pesquisadores acreditavam que esse fármaco se prendesse à topoisomerase e bloqueasse a ação de outras enzimas necessárias à síntese do DNA. Entretanto, uma pesquisa recente indica que ele modifica a topoisomerase de modo a incapacitar a remoção dos superenrolamentos à frente da replicação. O acúmulo de superenrolamentos pausa o maquinário de replicação e evita a proliferação das células de câncer. Como muitos outros fármacos ativos contra o câncer, a camptotecina também inibe a replicação de células normais, não cancerosas, o que explica como a quimioterapia afeta a saúde dos pacientes.

Este capítulo estuda a replicação do DNA, o processo pelo qual uma célula duplica seu DNA antes da divisão. Começamos com o mecanismo básico de replicação que surgiu a partir da estrutura do DNA apresentada por Watson e Crick. Então examinamos vários modos diferentes de replicação, as exigências da replicação e a direção universal da síntese do DNA. Examinamos também as enzimas e proteínas que participam do processo. Finalmente, analisamos os detalhes moleculares da

recombinação, que está intimamente relacionada com a replicação e é essencial para a separação dos cromossomos homólogos, a produção de variação genética e o reparo do DNA.

12.1 As informações genéticas precisam ser copiadas com exatidão a cada divisão celular

Em um jogo escolar, uma mensagem verbal, como “O cão marrom de João fugiu de casa”, é sussurrada para uma criança que repete a mensagem para outra criança. A mensagem é transmitida de uma criança para outra no pátio escolar até retornar para a primeira criança. Inevitavelmente, a última criança retorna com uma mensagem bem diferente, como “João Marrom tem um porco que vive sob a varanda”. Quanto maior o número de crianças brincando, mais distorcida a mensagem fica. Esse jogo ilustra um princípio importante: sempre surgem erros quando a informação é copiada, e, quanto mais vezes ela é copiada, maior o potencial para erros.

Um organismo complexo e multicelular lida com um problema parecido ao das crianças no jogo: como transmitir as instruções genéticas com fidelidade cada vez que a célula se divide. A solução para esse problema é fundamental para a replicação. Um único zigoto humano tem 6,4 bilhões de pares de base de DNA; mesmo uma pequena taxa de erro durante a cópia, como uma vez para cada milhão de pares de base, resultaria em 6.400 erros cada vez que a célula é dividida – erros que seriam combinados a milhões de divisões celulares que ocorrem no desenvolvimento humano.

O processo de cópia do DNA precisa ser não apenas muito exato, mas também muito rápido. O único cromossomo circular da *E. coli* tem cerca de 4,6 milhões de pares de base. Em uma taxa de mais de 1.000 nucleotídios por minuto, a replicação do cromossomo inteiro levaria quase 3 dias. Ainda, como já foi afirmado, essas bactérias são capazes de se dividir a cada 20 min. Na verdade, *Escherichia coli* replica seu DNA em uma taxa de 1.000 nucleotídios por segundo, com menos de um erro em um bilhão de nucleotídios. Como ocorre esse processo extraordinariamente exato e rápido?

12.2 Toda replicação de DNA ocorre de modo semiconservativo

A partir da estrutura tridimensional do DNA proposta por Watson e Crick em 1953 (ver Figura 10.7), várias implicações genéticas importantes estavam imediatamente aparentes. A natureza complementar das duas fitas de nucleotídios em uma molécula de DNA sugere que, durante a replicação, cada fita serve como molde para a síntese de uma nova fita. A especificidade do pareamento de bases (adenina com timina, guanina com citosina) indica que apenas uma sequência de bases pode ser especificada para cada molde, e então as duas moléculas de DNA construídas a partir do par de moldes serão idênticas à molécula original. Esse processo é chamado de **replicação semiconservativa**, porque cada fita original de nucleotídios permanece intacta (conservada), apesar de não se combinar mais com a mesma molécula; a molécula original de DNA tem sua metade (semi) conservada durante a replicação.

Inicialmente, foram propostos três modelos para a replicação do DNA. Na replicação conservativa (**Figura 12.1 A**), a molécula inteira de DNA de fita dupla serve como molde para uma nova molécula inteira de DNA, e a molécula de DNA original é *totalmente* conservada durante a replicação. Na replicação dispersiva (**Figura 12.1 B**), as duas fitas de nucleotídios se rompem (dispersam) em fragmentos, que servem como molde para a síntese de novos fragmentos de DNA, e então de alguma maneira se reúnem de novo em duas moléculas de DNA completas. Nesse modelo, cada molécula de DNA resultante é intercalada com fragmentos de DNA velho e novo; nenhuma molécula original é conservada. A replicação semiconservativa (**Figura 12.1 C**) é intermediária entre esses dois modelos; as duas fitas de nucleotídios se desenrolam, e cada uma serve como molde para uma nova molécula de DNA.

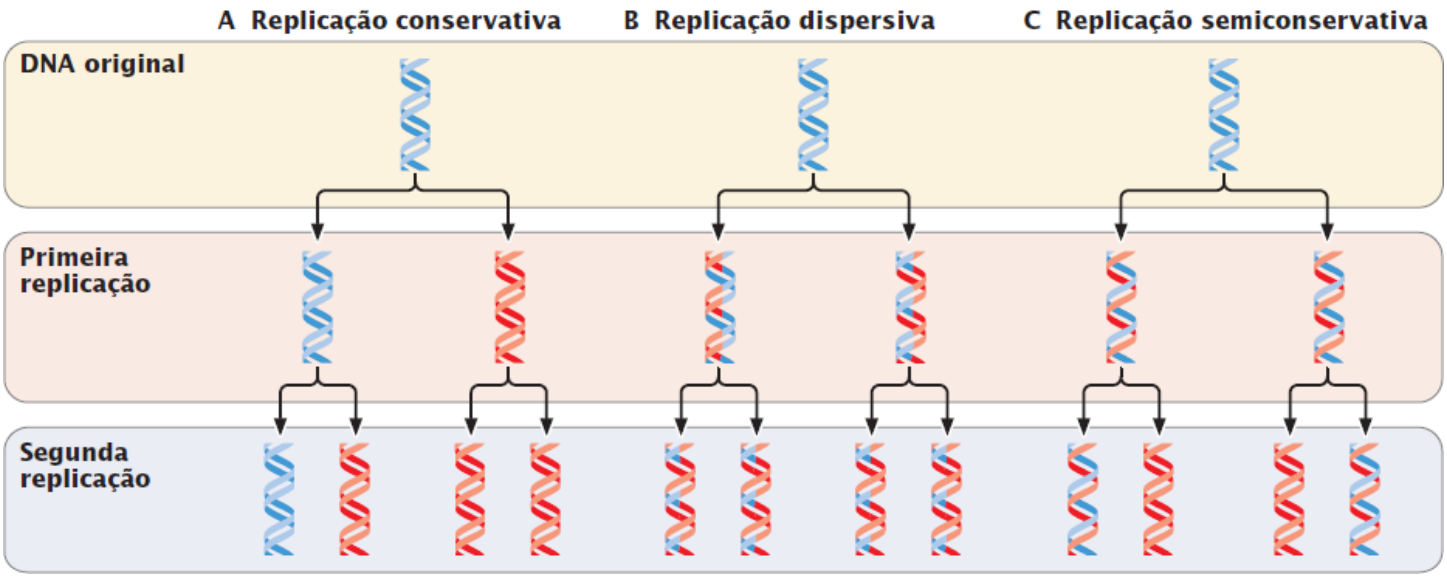


Figura 12.1 Os três modelos de replicação propostos são a replicação conservativa, a replicação dispersiva e a replicação semiconservativa.

Esses três modelos tornam possível diferentes predições sobre a distribuição do DNA original e do DNA recém-sintetizado após a replicação. Com a replicação conservativa, após um ciclo de replicação, 50% das moléculas seriam o DNA original e 50% seriam o DNA novo. Após um segundo ciclo de replicação, 25% das moléculas consistiriam totalmente em DNA original e 75% seriam totalmente constituídos por DNA novo. Com um novo adicional de replicação a proporção de moléculas com DNA novo aumentaria, embora o número de molécula com o DNA original permanecesse constante. A replicação dispersiva sempre produz moléculas híbridas, contendo DNA original e DNA novo, mas a proporção de DNA novo nas moléculas aumentaria após cada replicação. Em contrapartida, com a replicação semiconservativa, um ciclo de replicação produziria duas moléculas híbridas, cada uma com metade de DNA original e metade de DNA novo. Após um segundo ciclo de replicação, metade das moléculas seria híbrida e a outra metade seria apenas DNA novo. Ciclos adicionais de replicação produziriam mais e mais moléculas compostas inteiramente por DNA novo, e algumas moléculas híbridas persistiriam.

Experimento de Meselson e Stahl

Para determinar qual dos três modelos de replicação se aplica às células da *E. coli*, Matthew Meselson e Franklin Stahl precisavam diferenciar o DNA velho do DNA novo. Eles conseguiram usando dois isótopos de nitrogênio, ^{14}N (a forma comum) e ^{15}N (a forma rara, pesada). Meselson e Stahl cultivaram uma cultura de *E. coli* em um meio com ^{15}N como única fonte de nitrogênio; após várias gerações, todas as células da *E. coli* tinham ^{15}N incorporado em todas as bases purina e pirimidina de seu DNA (ver Figura 10.10). Os cientistas tiraram uma amostra dessas bactérias, trocaram o resto das bactérias para um meio com apenas ^{14}N e então retiraram outras amostras de bactérias durante as gerações seguintes. Em cada meio, o DNA bacteriano sintetizado antes da troca do meio continha ^{15}N e era relativamente pesado, enquanto qualquer DNA sintetizado após a troca do meio tinha ^{14}N e era relativamente leve.

Meselson e Stahl distinguiram entre o DNA marcado com ^{15}N pesado e o DNA com ^{14}N leve com o uso da **centrifugação por gradiente de densidade de equilíbrio** (**Figura 12.2**). Nessa técnica, um tubo de centrífuga é preenchido com uma solução de sal pesado e uma substância de densidade desconhecida – nesse caso, os fragmentos de DNA. O tubo é então centrifugado em uma centrífuga em altas velocidades. Após vários dias de centrifugação, um gradiente de densidade se desenvolve dentro do tubo, com a alta densidade no fundo e a baixa densidade no topo. A densidade dos fragmentos de DNA combina com a do sal: as moléculas leves sobem e as moléculas pesadas descem.

Meselson e Stahl descobriram que o DNA a partir das bactérias que crescem apenas em um meio com ^{15}N produzia uma única banda na posição esperada do DNA com apenas ^{15}N (**Figura 12.3 A**). O DNA das bactérias transferidas para o meio com ^{14}N e que sofreram um ciclo de replicação também produziu uma banda única, mas em uma posição intermediária entre o DNA esperado com apenas ^{15}N e o DNA esperado com apenas ^{14}N (**Figura 12.3 B**). Esse resultado não é consistente com o modelo de replicação conservativa, que prevê uma banda pesada (as moléculas originais de DNA) e uma banda leve (as novas moléculas de DNA). Uma única banda de densidade intermediária está prevista para os modelos semiconservativo e dispersivo.

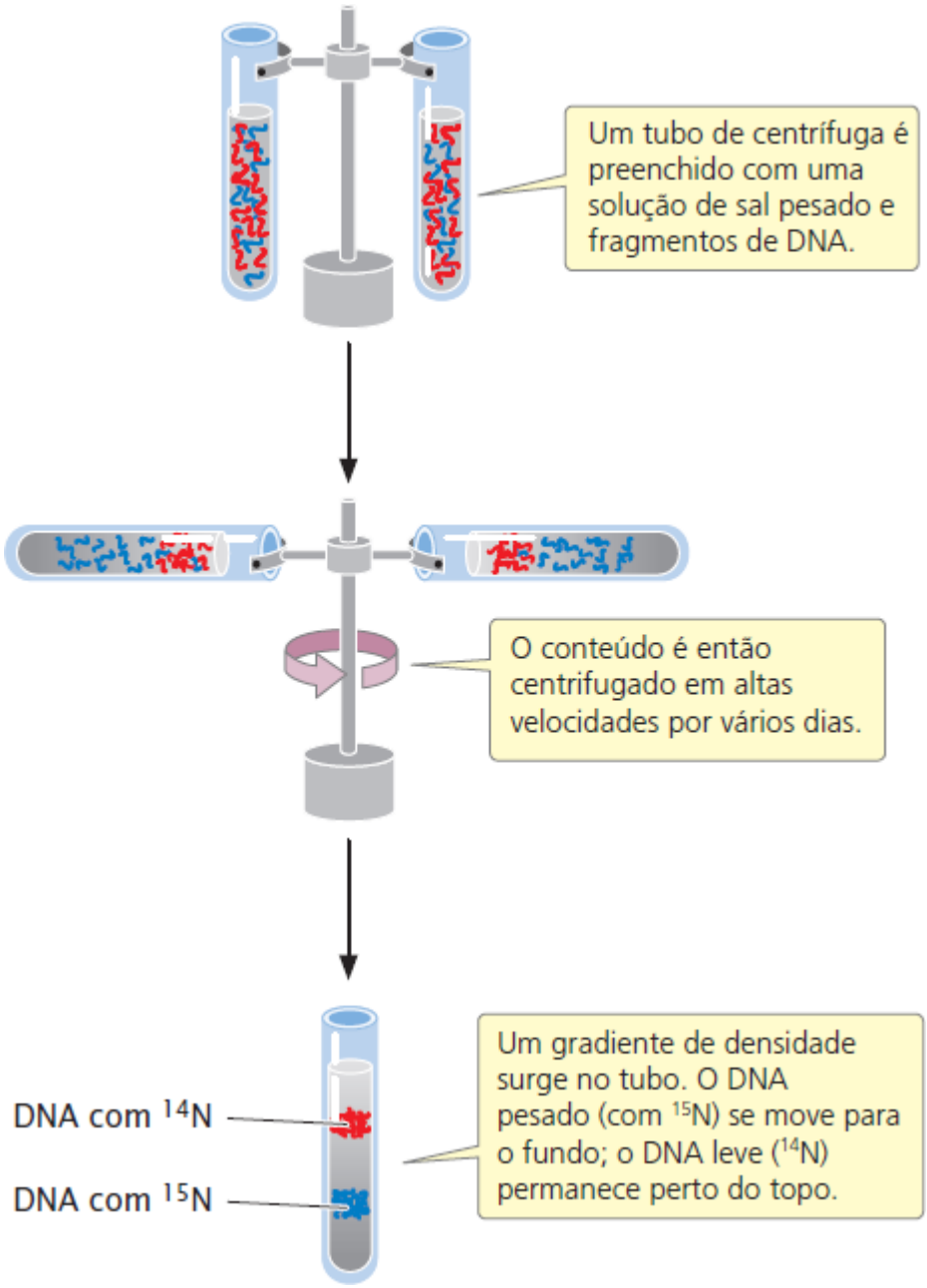


Figura 12.2 Meselson e Stahl usaram a centrifugação por gradiente de densidade de equilíbrio para diferenciar o DNA pesado, marcado com ^{15}N, do DNA leve, marcado com ^{14}N.

Para diferenciar entre esses dois modelos, Meselson e Stahl cultivaram as bactérias em um meio com ^{14}N por uma segunda geração. Após um segundo ciclo de replicação no meio com ^{14}N, apareceram duas bandas de mesma intensidade, uma na posição intermediária e outra na posição esperada para o DNA com ^{14}N (**Figura 12.3 C**). Todas as amostras retiradas após ciclos adicionais de replicação produziam as mesmas duas bandas, e a banda que representava o DNA leve se tornava progressivamente mais forte (**Figura 12.3 D**). Os resultados de Meselson e Stahl eram exatamente os resultados esperados para a replicação semiconservativa e eram incompatíveis com os resultados previstos para as replicações conservativa e dispersiva.

Resolva o Problema 22

Conceitos

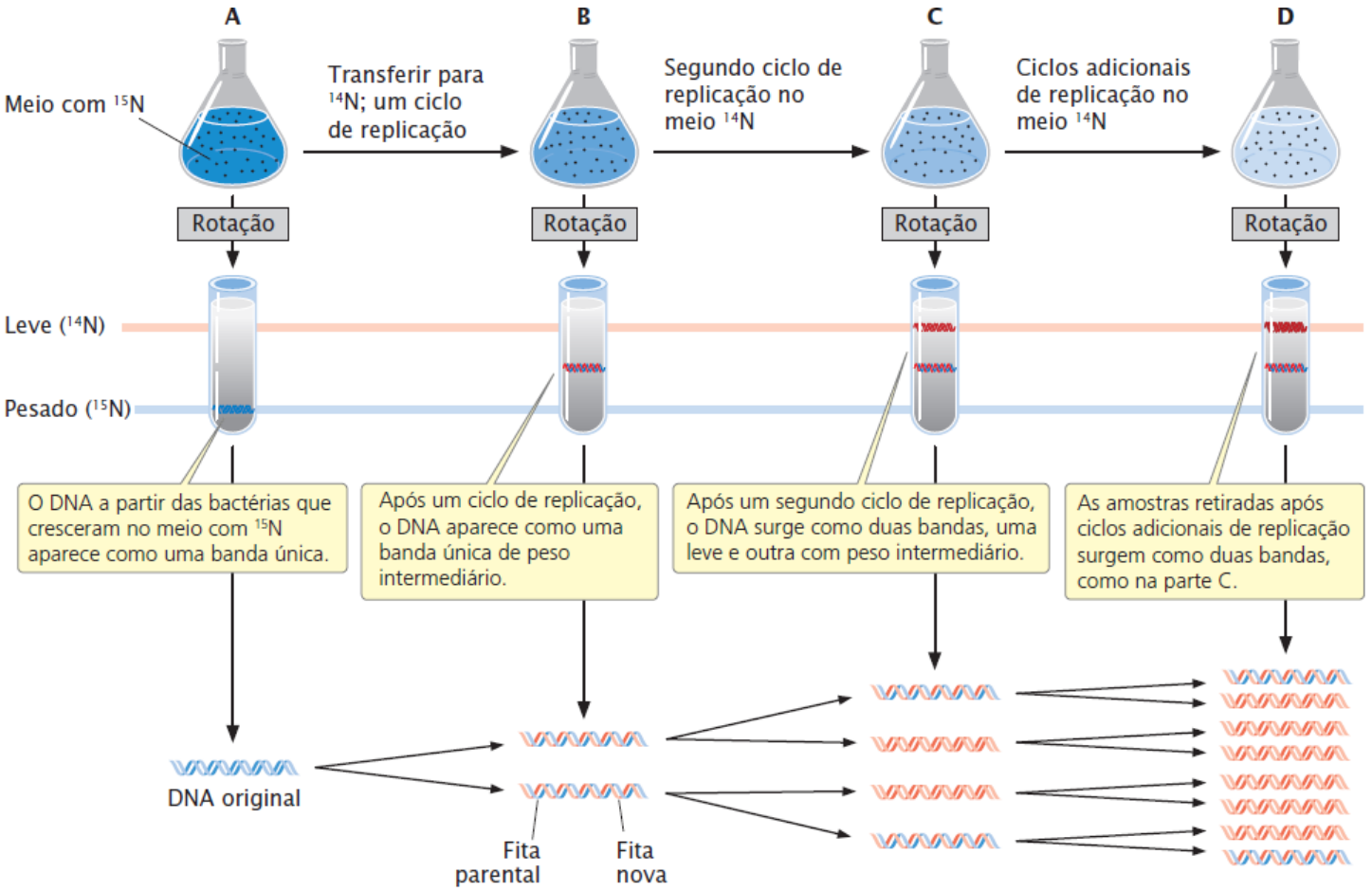
A replicação é semiconservativa: cada fita de DNA serve como molde para a síntese de uma nova molécula de DNA. Meselson e Stahl demonstraram de modo convincente que a replicação na *E. coli* é semiconservativa.

Checagem dos conceitos 1

Quantas bandas de DNA seriam esperadas no experimento de Meselson e Stahl após dois ciclos de replicação *conservativa*?

Experimento

Pergunta: Qual modelo de replicação de DNA – conservativa, dispersiva ou semiconservativa – aplica-se à *E. coli*?



Conclusão: A replicação do DNA da *E. coli* é semiconservativa.

Figura 12.3 Meselson e Stahl demonstraram que a replicação do DNA é semiconservativa.

Modos de replicação

Após o trabalho de Meselson e Stahl, os investigadores confirmaram que outros organismos também usam a replicação semiconservativa. Não foi encontrado indício da replicação conservativa ou dispersiva. Existem, entretanto, muitas formas diferentes nas quais a replicação semiconservativa pode ocorrer, com diferença principalmente na natureza do DNA molde, ou seja, se ele é linear ou circular.

As unidades de replicação são chamadas de **replicons**, cada uma com uma **origem de replicação**. A replicação inicia na origem e continua até o replicon inteiro ser replicado. Os cromossomos bacterianos têm uma única origem de replicação, enquanto os cromossomos eucarióticos têm várias origens.

Replicação teta. Um tipo comum de replicação que ocorre no DNA circular, como o encontrado na *E. coli* e em outras bactérias, é chamado de **replicação teta** (**Figura 12.4 A**), porque ela gera uma estrutura que lembra a letra grega teta (θ). Nesta figura e em todas as outras neste capítulo, a fita original (molde) de DNA é apresentada em cinza e a fita de DNA recém-sintetizada é apresentada em vermelho.

Na replicação teta, o DNA de fita dupla começa a se desenrolar na origem da replicação, produzindo fitas de nucleotídios de fita única que servem como moldes nos quais o novo DNA pode ser sintetizado. O desenrolamento da dupla-hélice gera uma alça, chamada **bolha de replicação**. O desenrolamento pode ser em uma ou em ambas as extremidades da bolha, tornando-a progressivamente maior. A replicação do DNA em ambas as fitas molde é simultânea com o desenrolamento. O ponto de desenrolamento, onde as duas fitas únicas de nucleotídios se separam da hélice de fita dupla do DNA, é chamado de **forquilha de replicação**.

Se existem duas forquilhas de replicação, uma em cada extremidade da bolha de replicação, as forquilhas continuam para fora nos dois sentidos em um processo chamado de **replicação bidirecional**, simultaneamente abrindo e replicando o DNA até elas se encontrarem. Se houver uma única forquilha de replicação, ela prossegue ao redor do círculo inteiro. Tanto a replicação bidirecional quanto a unidirecional produzem duas moléculas de DNA circulares completas, cada uma com uma fita antiga e uma fita nova de nucleotídios.

John Cairns forneceu a primeira evidência visível da replicação teta em 1963 ao cultivar bactérias na presença de nucleotídios radioativos. Após a replicação, cada molécula de DNA tinha uma fita "quente" (radioativa) e uma fita "fria" (não radioativa). Cairns isolou o DNA das bactérias após a replicação, colocou em uma grelha de microscopia eletrônica e então a cobriu com uma emulsão fotográfica. A radioatividade na amostra expôs a emulsão e produziu uma imagem da molécula (chamada de autorradiografia) semelhante à maneira pela qual a luz expõe um filme fotográfico. Como o DNA recém-sintetizado tem nucleotídios radioativos, Cairns pode produzir uma micrografia eletrônica do processo de replicação, semelhante ao processo apresentado na **Figura 12.4 B**.

Replicação por círculo rolante. Outro modo de replicação, chamado **replicação por círculo rolante** (**Figura 12.5**), ocorre em alguns vírus e no fator F (um pequeno círculo de DNA extracromossômico que controla a reprodução, discutido no *Capítulo 9*) da *E. coli*. Esse modo de replicação é iniciado por uma ruptura em uma das fitas de nucleotídios que cria um grupo 3′-OH e um grupo 5′-fosfato. Novos nucleotídios são adicionados à extremidade 3′ da fita rompida, com a fita interna (não rompida) servindo como molde. À medida que novos nucleotídios são adicionados à extremidade 3′, a extremidade 5′ da fita rompida é deslocada do molde, rolando como um fio sendo retirado de um carretel. A extremidade 3′ cresce ao redor do círculo, originando o nome de círculo rolante.

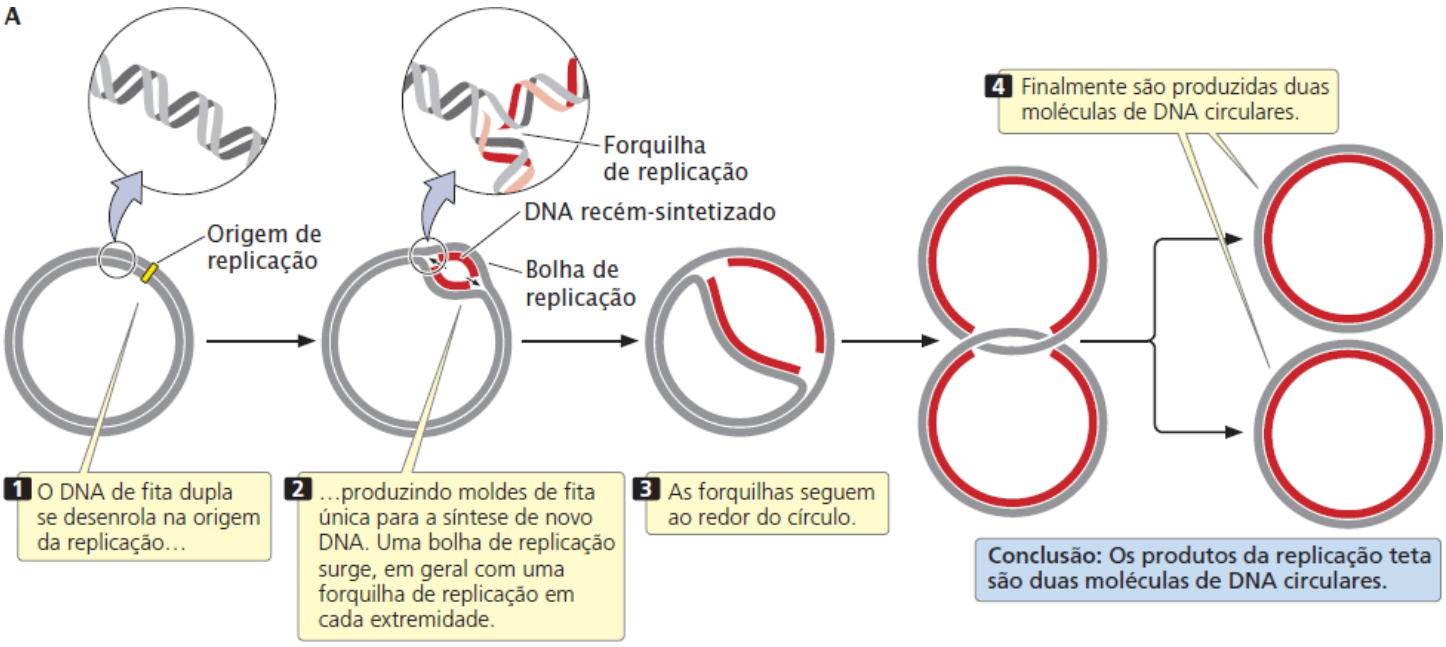


Figura 12.4 A replicação teta é um tipo de replicação comum na *E. coli* e em outros organismos com DNA circular. (*Parte B*: Bernhard Hirt, L'Institut Suisse de Recherche Éxperimentale sur le Cancer.)

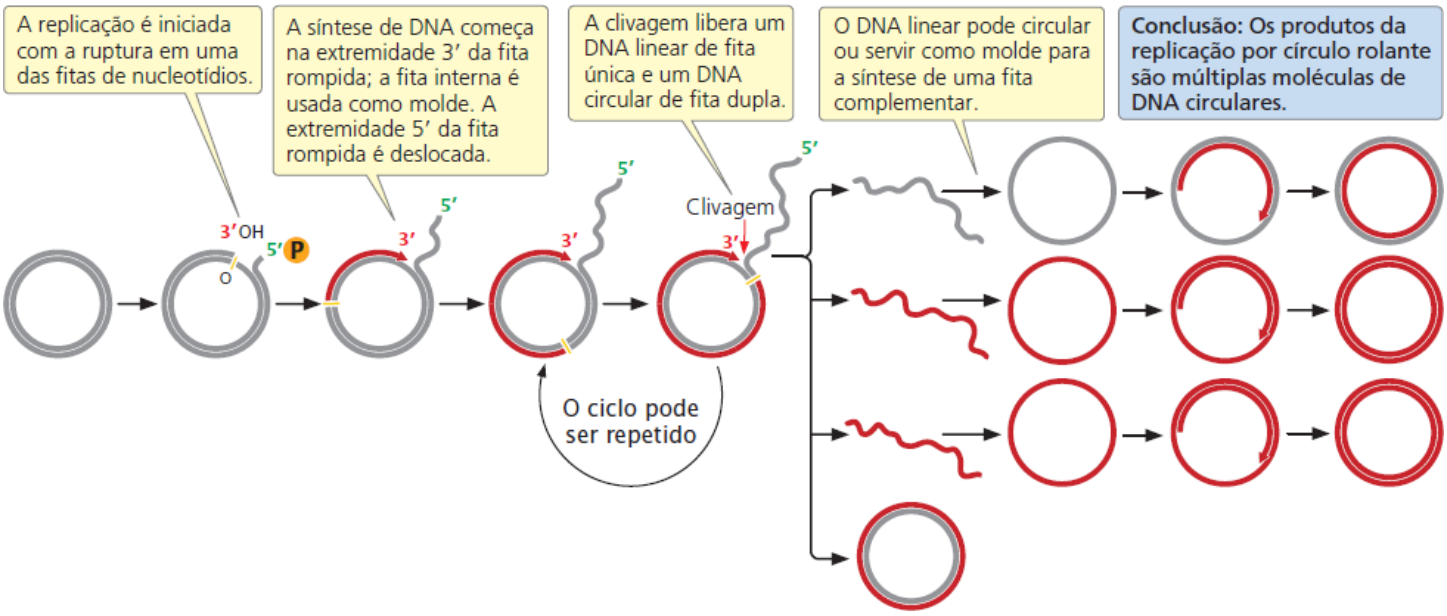


Figura 12.5 **A replicação por círculo rolante ocorre em alguns vírus e no fator F da *E. coli*.**

A forquilha de replicação continua ao redor do círculo várias vezes, produzindo várias cópias ligadas da mesma sequência. Com cada volta ao redor do círculo, a extremidade 3′ crescente desloca a fita de nucleotídios sintetizada na volta seguinte. Finalmente, a molécula de DNA linear é rompida do círculo, gerando uma molécula de DNA circular de fita dupla e uma molécula de DNA linear de fita única. A molécula linear circula antes ou após servir como molde para a síntese de uma fita complementar.

Replicação eucariótica linear. As moléculas de DNA circulares que sofrem replicação teta ou por círculo rolante têm uma única origem de replicação. Por causa do tamanho limitado dessas moléculas de DNA, a replicação que inicia a partir da origem consegue atravessar o cromossomo inteiro em um tempo razoável. Os cromossomos lineares maiores nas células eucarióticas, entretanto, têm muito DNA para ser replicado rapidamente a partir de uma única origem. A replicação eucariótica prossegue em uma taxa de 500 a 5.000 nucleotídios por minuto em cada forquilha de replicação (consideravelmente mais lenta que a replicação bacteriana). Mesmo a 5.000 nucleotídios por minuto em cada forquilha, a síntese do DNA iniciando a partir de uma única origem levaria 7 dias para replicar um típico cromossomo humano com 100 milhões de pares de base de DNA. A replicação dos cromossomos eucarióticos ocorre em questão de minutos ou horas, e não dias. Essa velocidade é possível porque a replicação inicia em milhares de origens.

Os replicons eucarióticos típicos têm de 20.000 a 300.000 pares de base de comprimento (**Quadro 12.1**). Em cada origem de replicação, o DNA se desenrola e produz uma bolha de replicação. A replicação ocorre em ambas as fitas em cada extremidade da bolha, com as duas forquilhas de replicação se espalhando para fora. Por fim, as forquilhas de replicação dos replicons adjacentes correm na direção uma da outra, e os replicons se unem para formar longos trechos de DNA recém-sintetizado (**Figura 12.6**). A replicação e a fusão de todos os replicons geram duas moléculas idênticas de DNA. As características importantes das replicações teta, por círculo rolante e eucariótica linear, estão resumidas no **Quadro 12.2**. Resolva o Problema 23

Conceitos

A replicação teta, a replicação por círculo rolante e a replicação linear apresentam diferenças com relação ao início e progresso da replicação, mas todas produzem novas moléculas de DNA por meio da replicação semiconservativa.

Checagem dos conceitos 2

Que tipo de replicação requer a ruptura na fita de nucleotídios para começar?

a. Replicação teta.
b. Replicação por círculo rolante.
c. Replicação eucariótica linear.
d. Todas as opções anteriores.

Quadro 12.1 Número e tamanho dos replicons.

Organismo	Número de origens de replicação	Tamanho médio do replicon (pb)
Escherichia coli (bactéria)	1	4.600.000
Saccharomyces cerevisiae (levedura)	500	40.000
Drosophila melanogaster (mosca-da-fruta)	3.500	40.000
Xenopus laevis (sapo)	15.000	200.000
Mus musculus (camundongo doméstico)	25.000	150.000

Fonte: Dados de B. L. Lewin, Genes V (Oxford: Oxford University Press, 1994), p. 536.

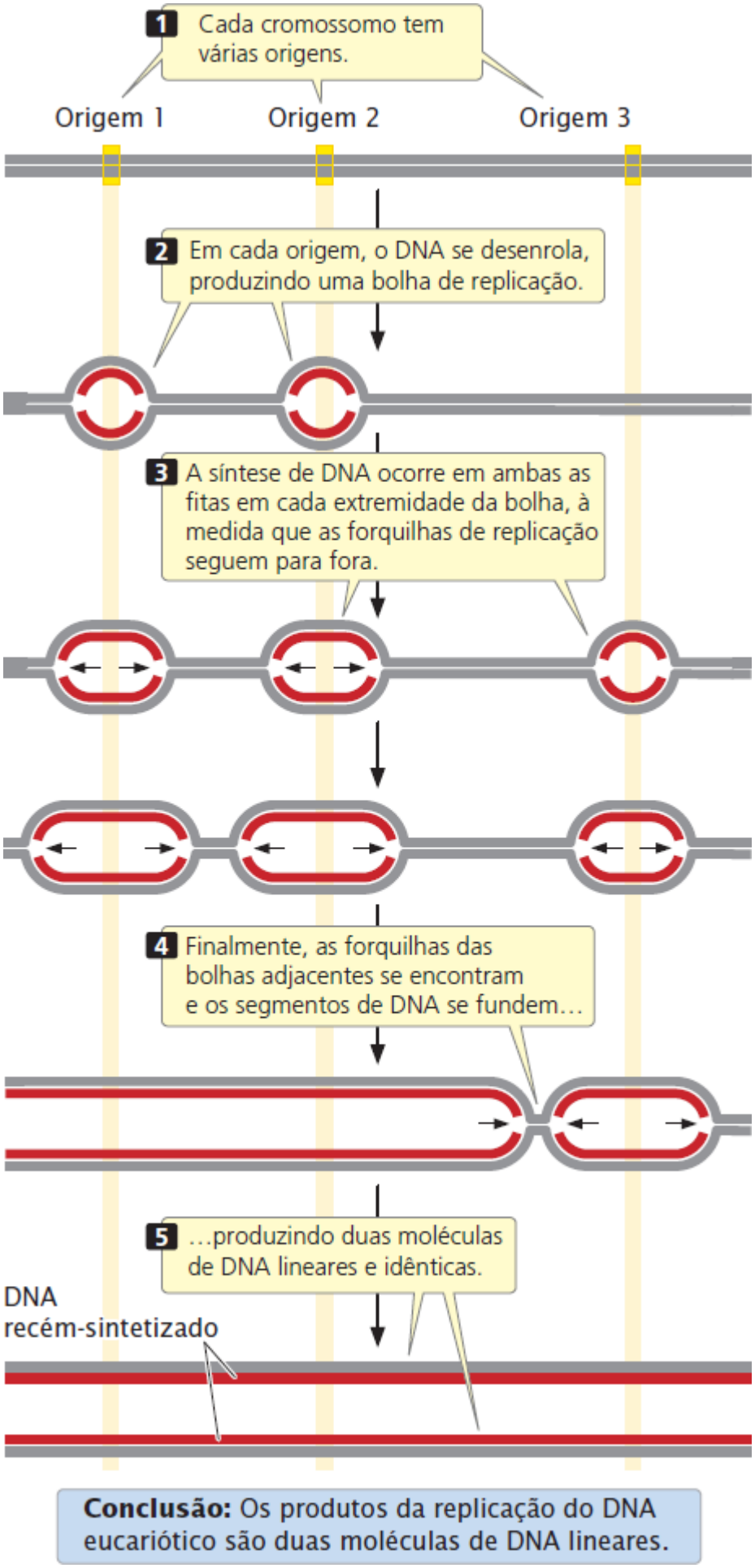


Figura 12.6 **A replicação de DNA linear ocorre nos cromossomos eucarióticos.**

Quadro 12.2 Características da replicação teta, por círculo rolante e eucariótica linear.

Modelo de replicação	Molde de DNA	Ruptura da fita de nucleotídios	Número de replicons	Unidirecional ou bidirecional	Produtos
Teta	Circular	Não	1	Unidirecional ou bidirecional	Duas moléculas circulares
Por círculo rolante	Circular	Sim	1	Unidirecional	Uma molécula circular e uma molécula linear que pode circular
Eucariótica linear	Linear	Não	Muitas	Bidirecional	Duas moléculas lineares

Exigências da replicação

Embora o processo de replicação inclua muitos componentes, eles podem ser combinados em três grupos principais:

1. Um molde de DNA de fita dupla.
2. Matérias-primas (substratos) a serem montadas em uma nova fita de nucleotídios.
3. Enzimas e outras proteínas que “leem” o molde e reúnem os substratos em uma molécula de DNA.

Graças à natureza semiconservativa da replicação do DNA, uma molécula de DNA de fita dupla precisa se desenvolver para expor as bases que atuam como molde para a montagem de novas fitas de polinucleotídios, que serão complementares e antiparalelas às fitas molde. As matérias-primas a partir das quais novas moléculas de DNA são sintetizadas são os trifosfatos de desoxirribonucleosídio (dNTPs), cada um composto por um açúcar desoxirribose e uma base (um nucleosídio) ligados a três grupos fosfato (**Figura 12.7 A**). Na síntese de DNA, os nucleotídios são acondicionados ao grupo 3′-OH da fita de nucleotídios crescente (**Figura 12.7 B**). O grupo 3′-OH do último nucleotídio na fita fixa-se ao grupo 5′-fosfato do dNTP de chegada. Os dois grupos fosfato são rompidos a partir do dNTP de chegada, e uma ligação fosfodiéster é criada entre os dois nucleotídios.

A síntese de DNA não acontece espontaneamente. Ela requer um grupo de enzimas e proteínas que atuam de maneira coordenada. Examinaremos esse complexo arranjo de proteínas e enzimas à medida que analisarmos o processo de replicação com mais detalhes.

Conceitos

A síntese de DNA requer um molde de DNA de fita única, trifosfatos de desoxirribonucleosídio, uma fita de nucleotídios crescente e um grupo de enzimas e proteínas.

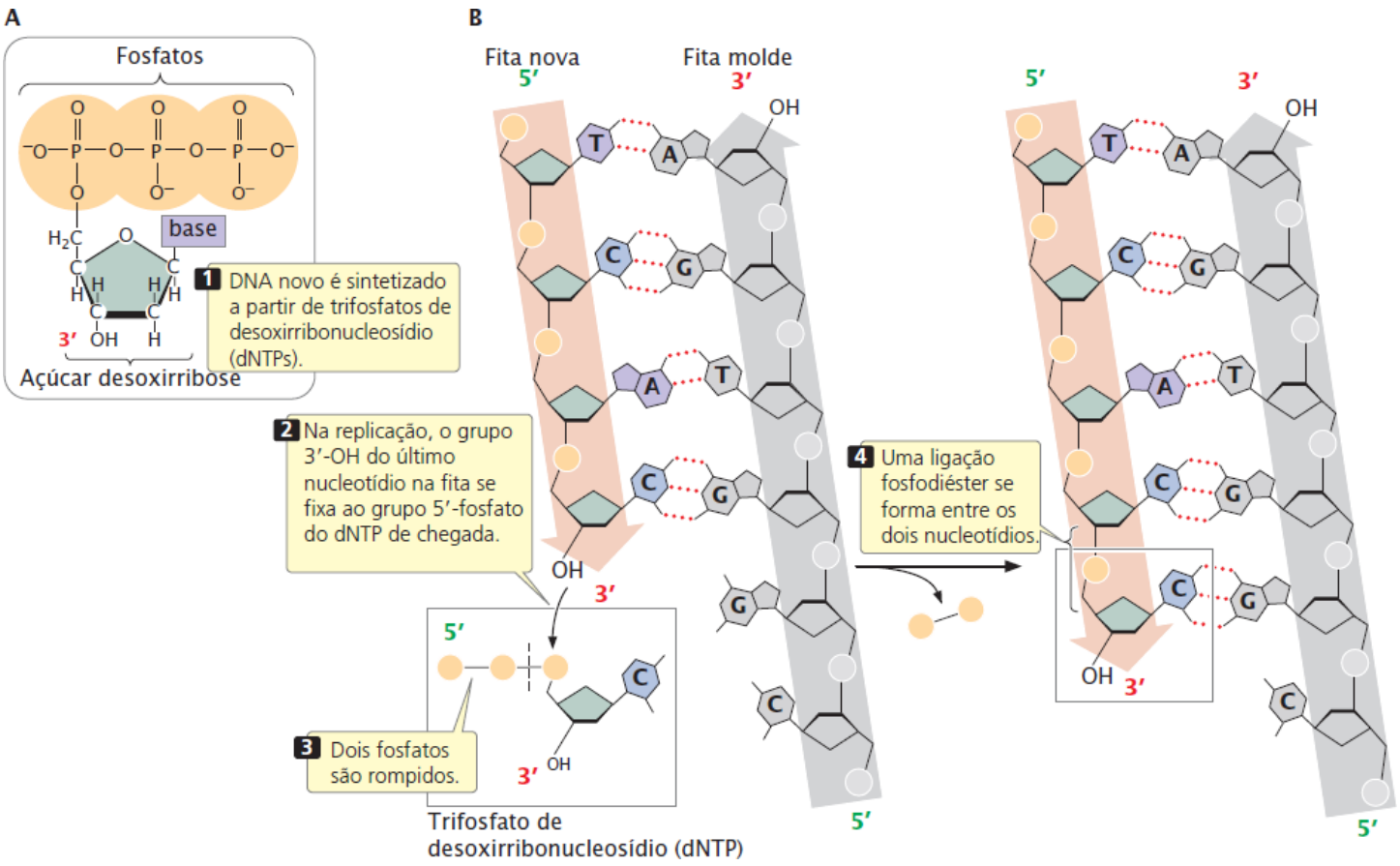


Figura 12.7 O novo DNA é sintetizado a partir de trifosfatos de desoxirribonucleosídio (dNTPs). A fita recém-sintetizada é complementar e antiparalela à fita molde; as duas fitas estão mantidas unidas por pontes de hidrogênio (representadas por linhas vermelhas pontilhadas) entre as bases.

Sentido de replicação

Na síntese do DNA, novos nucleotídios são unidos, um de cada vez, à extremidade 3′ da fita recém-sintetizada. As **DNA polimerases**, enzimas que sintetizam DNA, podem adicionar nucleotídios apenas na extremidade 3′ da fita crescente (não na extremidade 5′), e então novas fitas de DNA sempre alongam no mesmo sentido 5′ para 3′ (5′ → 3′). Como os dois moldes de DNA de fita única são antiparalelos e o prolongamento de fita é sempre 5′ → 3′, se a síntese em um molde segue, por exemplo, da direita para a esquerda, então a síntese do outro molde deve seguir no sentido oposto, da esquerda para a direita (**Figura 12.8**). À medida que o DNA se desenrola durante a replicação, a natureza antiparalela das duas fitas de DNA significa que um molde é exposto no sentido 5′ → 3′ e o outro é exposto no sentido 3′ → 5′. Então, como a síntese pode ocorrer simultaneamente em ambas as fitas na forquilha?

Replicação contínua e descontínua. À medida que o DNA se desenrola, a fita molde exposta no sentido 3′ → 5′ (a fita inferior nas Figuras 12.8 e 12.9) possibilita a síntese contínua da nova fita (no sentido 5′ → 3′). Essa nova fita, que sofre **replicação contínua**, é denominada **fita líder**.

A outra fita molde é exposta no sentido 5′ → 3′ (a fita superior nas Figuras 12.8 e 12.9). À medida que a extensão curta do DNA for desenrolada, a síntese deve seguir 5′ → 3′; ou seja, no sentido *oposto* da fita não desenrolada (**Figura 12.9**). Como apenas uma extensão curta de DNA precisa estar não enrolada antes de a síntese na sua fita começar, o maquinário de replicação logo fica sem molde. Nesse momento, mais DNA está desenrolando, fornecendo um novo molde na extremidade 5′ da nova fita. A síntese de DNA deve iniciar novamente na forquilha de replicação e seguir na direção oposta do movimento da forquilha até chegar ao segmento previamente replicado de DNA. Esse processo é repetido várias vezes, então a síntese dessa fita ocorre em explosões curtas, descontínuas. A fita recém-criada que sofre **replicação descontínua** é chamada de **fita tardia**.

Fragmentos de Okazaki. As extensões curtas de DNA produzidas pela replicação descontínua da fita tardia são chamadas de **fragmentos de Okazaki**, descobertos por Reiji Okazaki. Nas células bacterianas, cada fragmento de Okazaki varia de 1.000 a 2.000 nucleotídios de extensão; nas células eucarióticas, varia de 100 a 200 nucleotídios. Esses fragmentos na fita tardia estão ligados para criar uma nova molécula de DNA contínua.

Conceitos

Toda síntese de DNA é 5′ →3′, o que significa que novos nucleotídios são sempre adicionados à extremidade 3′ da fita de nucleotídios crescente. Em cada forquilha de replicação, a síntese da fita líder segue de maneira contínua, e a da fita tardia segue de maneira descontínua.

Checagem dos conceitos 3

A replicação descontínua é o resultado de qual propriedade do DNA?

a. Bases complementares.
b. Grupo fosfato carregado.
c. Fitas de nucleotídios antiparalelas.
d. Açúcar de cinco carbonos.

Conceitos conectantes

O sentido da síntese em diferentes modelos de replicação

Vamos relacionar o sentido da síntese de DNA aos tipos de replicação examinados anteriormente. No modelo teta (**Figura 12.10 A**), o DNA se desenrola em um lugar particular, a origem, e uma bolha de replicação é formada. Se a bolha tem duas forquilhas, uma em cada extremidade, a síntese ocorre simultaneamente em ambas as forquilhas (replicação bidirecional). Em cada forquilha, a síntese em uma das fitas molde prossegue no mesmo sentido que a fita desenrolada; essa fita recém-replicada é a fita líder com replicação contínua. Na outra fita molde, a síntese segue no sentido oposto que o da fita desenrolada; essa fita recém-sintetizada é a fita tardia com replicação descontínua. Observe apenas uma das fitas molde dentro da bolha. Veja que a síntese nessa fita é contínua em uma forquilha, mas descontínua em outra. Essa diferença surge porque a síntese do DNA está sempre no mesmo sentido (5′ → 3′), mas as duas forquilhas se movem em sentidos opostos.

A replicação no modelo por círculo rolante (**Figura 12.10 B**) é diferente, porque não há bolha de replicação. A replicação começa na extremidade 3′ da fita de nucleotídios rompida. A replicação contínua ocorre no molde circular à medida que novos nucleotídios são adicionados à extremidade 3′.

A replicação de moléculas de DNA lineares, como as encontradas nas células eucarióticas, produz várias bolhas de replicação (**Figura 12.10 C**). A síntese de DNA nessas bolhas é a mesma da única bolha de replicação do modelo teta; ela começa no centro de cada bolha de replicação e segue em duas forquilhas, uma em cada extremidade da bolha. Em ambas as forquilhas, a síntese da fita líder segue no mesmo sentido que da fita desenrolada, enquanto a síntese da fita tardia segue no sentido oposto. Resolva o Problema 25 A a C

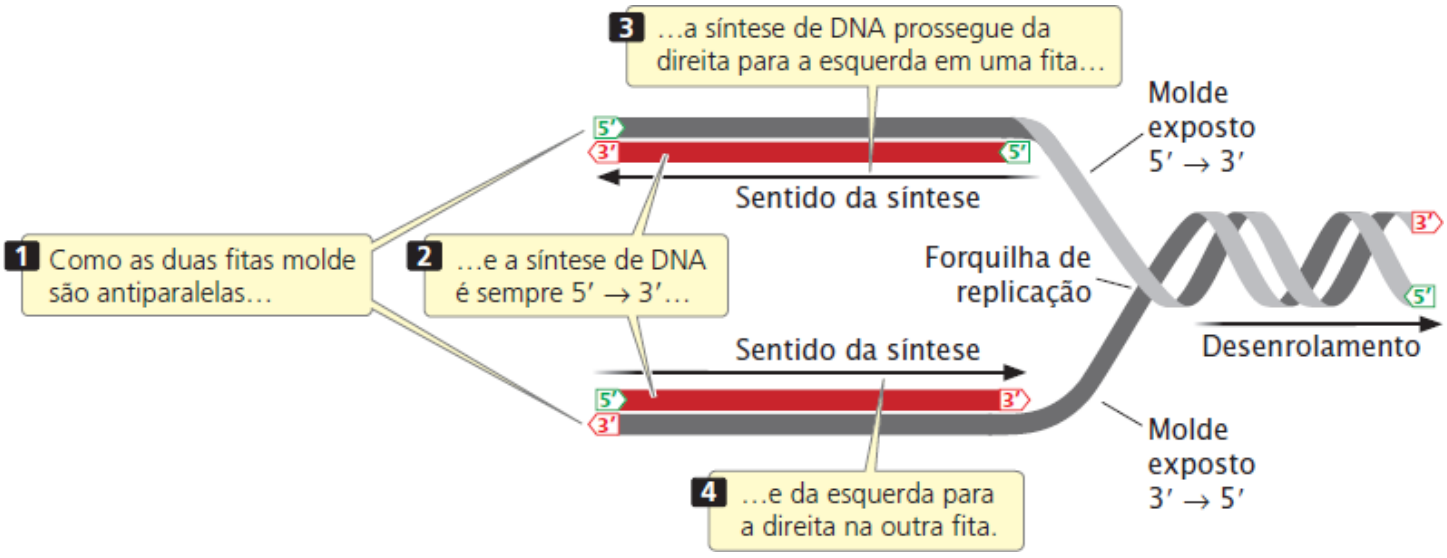


Figura 12.8 A síntese de DNA ocorre em sentidos opostos nas duas fitas molde de DNA. A replicação do DNA em uma única forquilha começa quando uma molécula de DNA de fita única se desenrola para fornecer dois moldes de fita única.

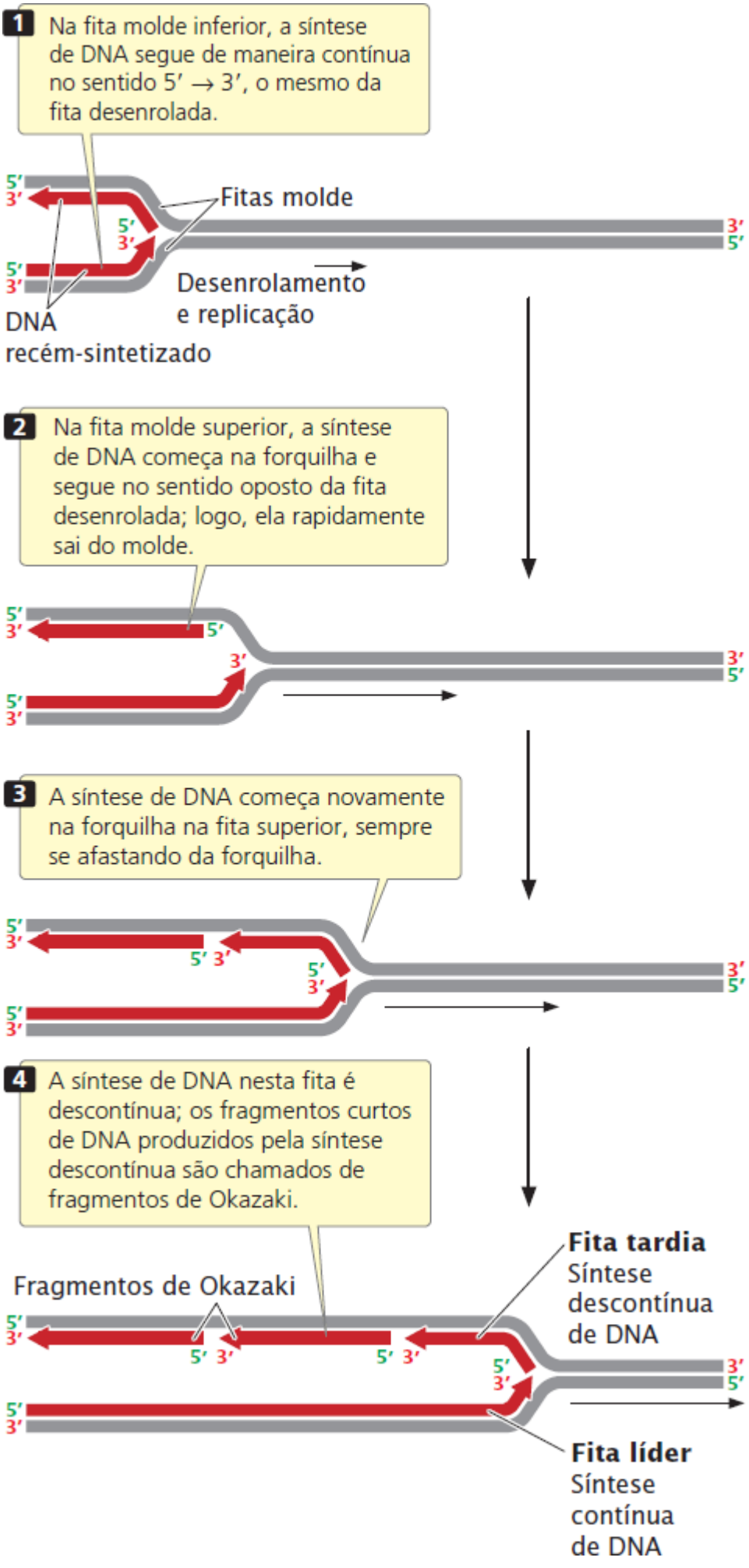


Figura 12.9 **A síntese do DNA é contínua em uma fita molde de DNA e descontínua em outra.**

12.3 A replicação bacteriana exige numerosas enzimas e proteínas

A replicação ocorre em quatro estágios: iniciação, desenrolamento, prolongamento e terminação. A discussão do processo de replicação a seguir se concentra nos sistemas bacterianos, nos quais a replicação foi estudada a fundo e é mais bem compreendida. Embora muitos aspectos da replicação nas células eucarióticas sejam semelhantes aos das células bacterianas, existem diferenças importantes. Vamos comparar a replicação bacteriana e eucariótica posteriormente neste capítulo.

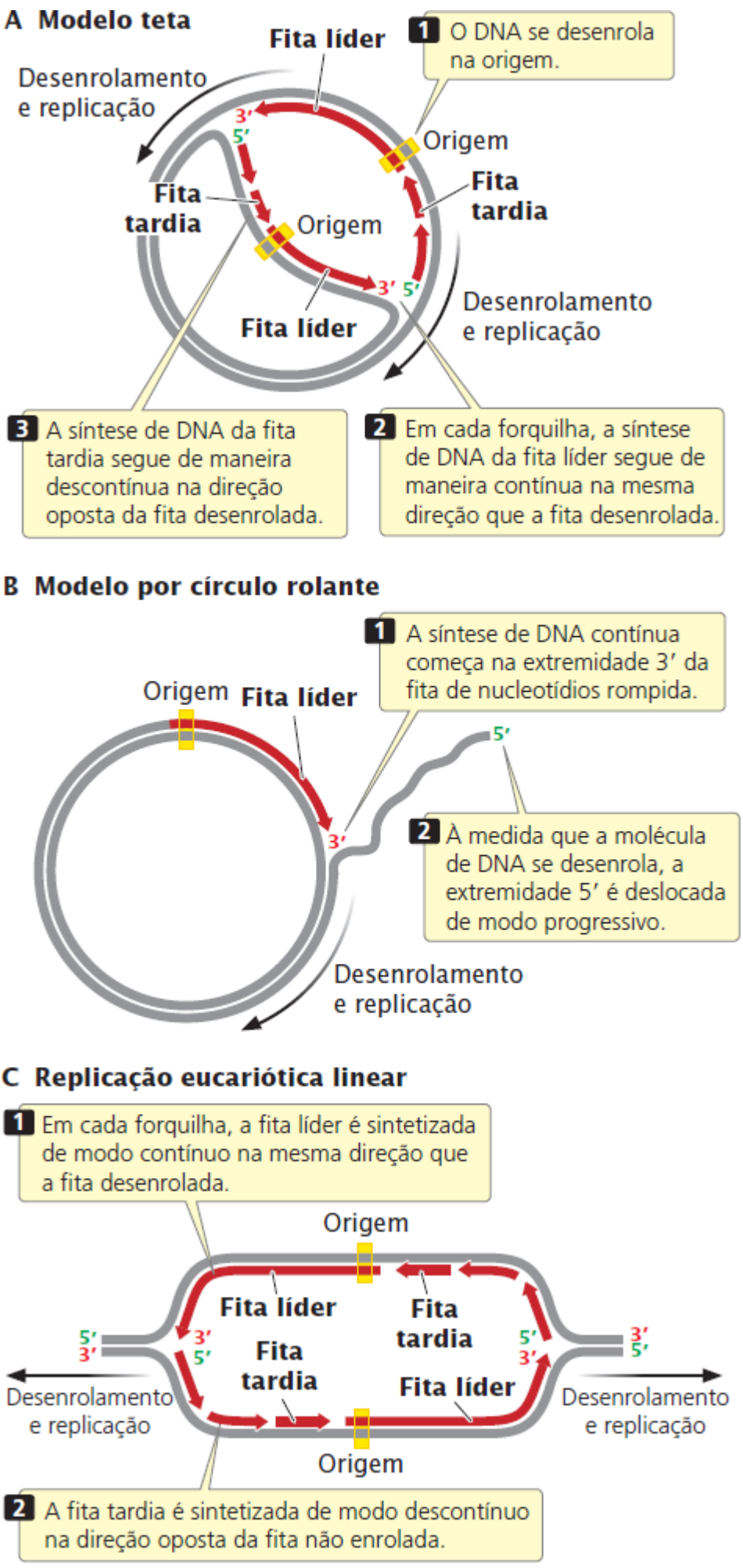


Figura 12.10 O processo de replicação é diferente na replicação teta, na replicação por círculo rolante e na replicação linear.

Iniciação

O cromossomo circular da *E. coli* tem uma única origem de replicação (*oriC*). A sequência mínima necessária para *oriC* funcionar é composta por 245 pb com vários sítios críticos. Uma **proteína de iniciação** (conhecida como DnaA na *E. coli*) liga-se à *oriC* e faz um segmento curto do DNA se desenrolar. Esse desenrolamento torna possível que a helicase e outras proteínas que se ligam à fita única prendam-se à fita de polinucleotídios (**Figura 12.11**).

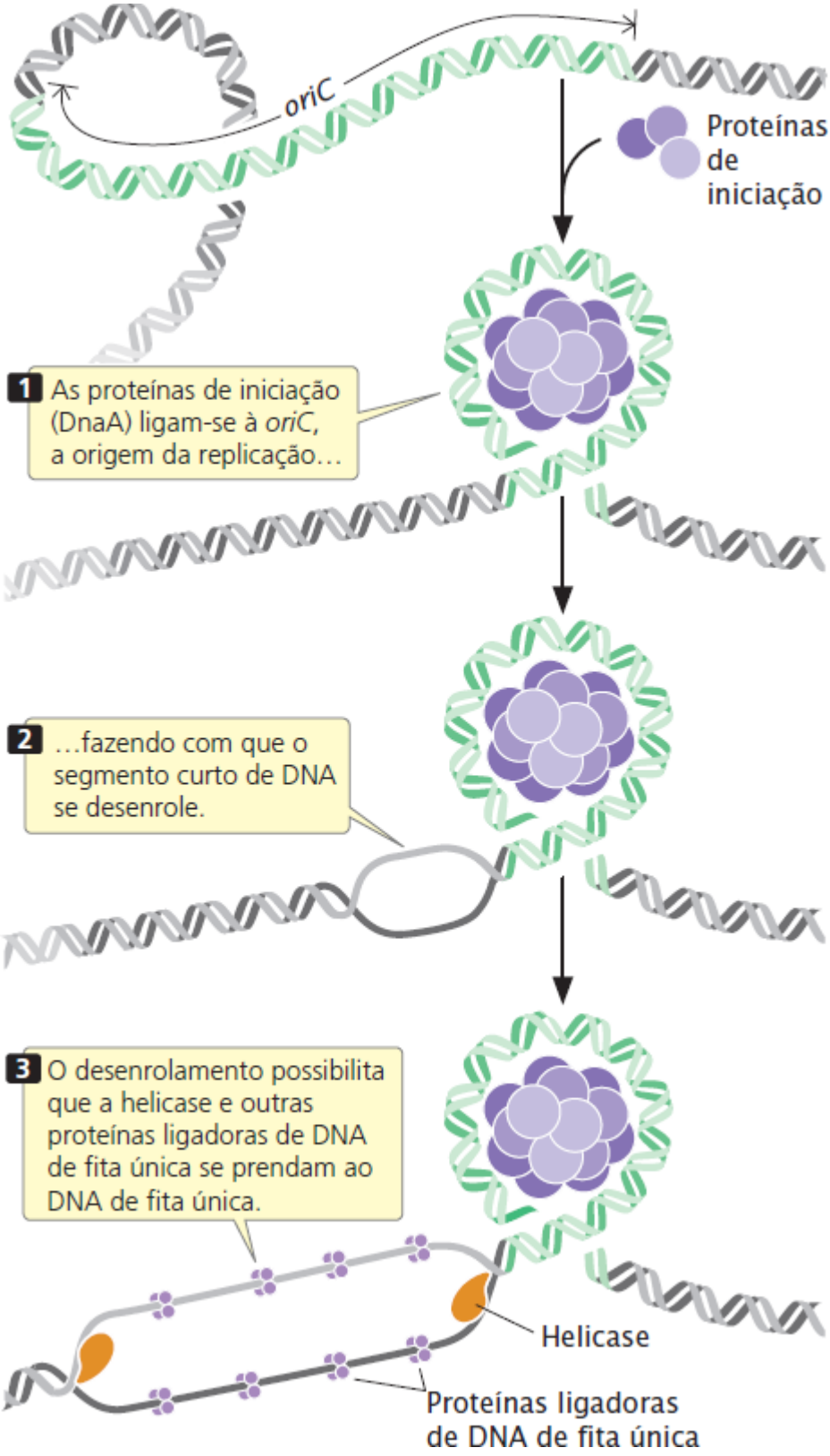


Figura 12.11 A replicação do DNA da *E. coli* começa quando as proteínas de iniciação se ligam à *oriC*, a origem da replicação.

Desenrolamento

Como a síntese do DNA requer um molde de fita única e o DNA de fita dupla tiver de ser desenrolado antes que ocorra a síntese de DNA, a célula depende de várias proteínas e enzimas para fazer a desenrolamento.

DNA helicase. A **DNA helicase** rompe as pontes de hidrogênio entre as bases de duas fitas de nucleotídios de uma molécula de DNA. A helicase não pode *iniciar* o desenrolamento do DNA de fita dupla; a proteína de iniciação separa primeiro as fitas de DNA na origem, fornecendo um curto segmento de DNA de fita dupla onde essa enzima se liga. A helicase liga-se ao molde da fita tardia em cada forquilha de replicação e move-se na direção 5′ →3′ ao longo dessa fita, também movendo a forquilha (**Figura 12.12**).

Proteínas ligadoras de DNA de fita única. Após o DNA ser desenrolado pela helicase, as **proteínas ligadoras de DNA de fita única** (SSBs) prendem-se firmemente ao DNA de fita única exposto (ver Figura 12.12). Essas proteínas protegem as cadeias de nucleotídios de fita única e evitam a formação de estruturas secundárias como os grampos (ver Figura 10.17) que interferem na replicação. De forma diversa à de muitas proteínas ligadoras de DNA, as SSBs são indiferentes à sequência de base: elas se ligam a qualquer DNA de fita única. Essas proteínas formam tetrâmeros (grupos de quatro); cada tetrâmero abrange de 35 a 65 nucleotídios.

DNA girase. Outra proteína essencial para o processo de desenrolamento é a enzima **DNA girase**, uma topoisomerase. Conforme discutido no Capítulo 11 e na introdução deste capítulo, as topoisomerases controlam o superenrolamento do DNA. Existem dois tipos principais dessa enzima: as topoisomerases tipo I alteram o superenrolamento ao fazer rupturas de fita única no DNA, enquanto as topoisomerases tipo II criam rupturas de fita dupla. A DNA girase é uma topoisomerase tipo II. Na replicação, ela reduz a tensão de torção (torque) que surge à frente da forquilha de replicação como resultado do desenrolamento (ver Figura 12.12). Ela reduz o torque ao romper a fita dupla em um segmento da hélice de DNA, passando outro segmento da hélice através da ruptura e, então, liberando as extremidades rompidas do DNA. Essa ação, que requer ATP, remove uma deformação no DNA e reduz o superenrolamento.

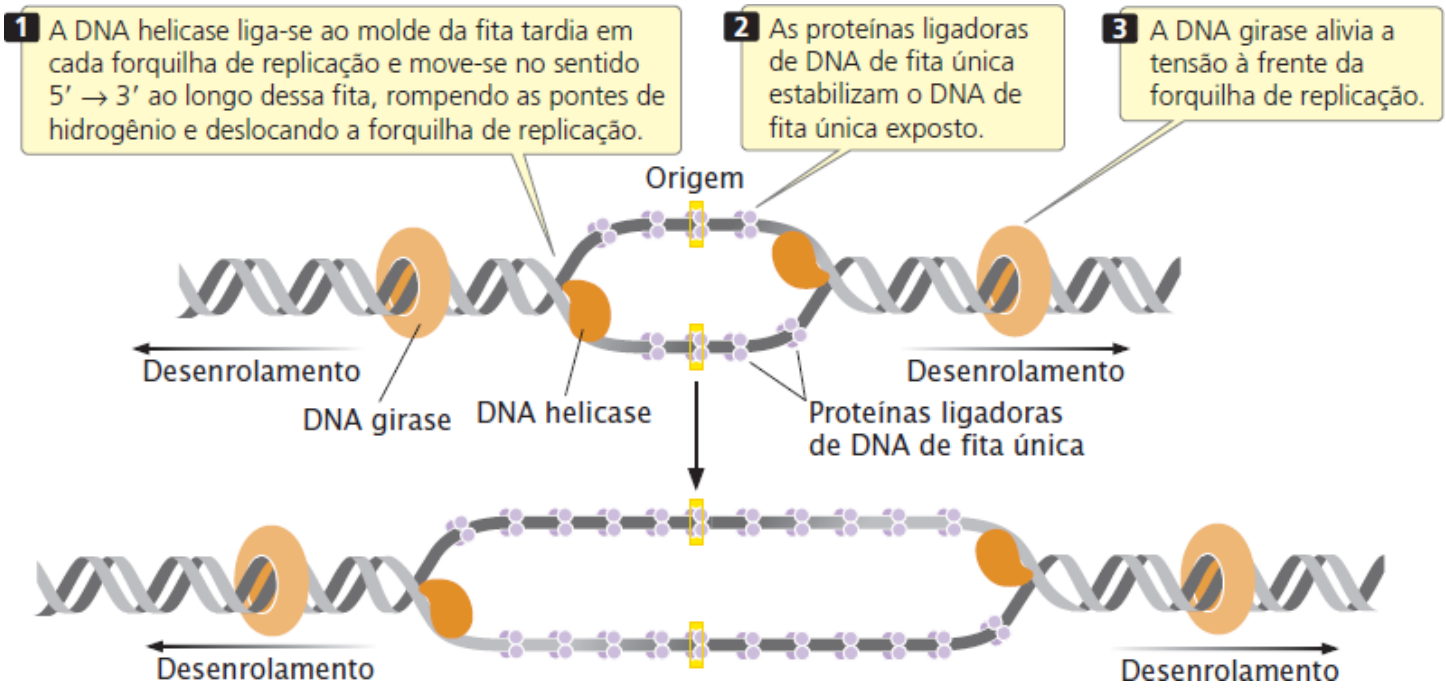


Figura 12.12 A DNA helicase desenrola o DNA ao se ligar ao molde de fita tardia em cada forquilha de replicação e mover-se no sentido 5′ → 3′.

Um grupo de antibióticos chamado de 4-quinolonas mata bactérias ao ligar-se à DNA girase e inibir sua ação. A inibição da DNA girase leva ao término da síntese de DNA e do crescimento bacteriano. Um exemplo de 4-quinolona é o ácido nalidíxico, introduzido pela primeira vez nos anos 1960 e que costumava ser usado para tratar infecções urinárias. Muitas bactérias adquiriram resistência às quinolonas por meio de mutações no gene para DNA girase.

Conceitos

A replicação é iniciada em uma origem de replicação, onde uma proteína de iniciação se liga e faz com que um pequeno segmento do DNA se desenrole. A enzima DNA helicase rompe as pontes de hidrogênio em uma forquilha de replicação e as proteínas ligadoras de DNA de fita única estabilizam as fitas separadas. A DNA girase reduz a tensão de torção que surge à medida que as duas fitas de DNA de dupla-hélice se

desenrolam.

Checagem dos conceitos 4

Ordene os seguintes componentes na sequência em que são usados durante a replicação: helicase, proteína ligadora de DNA de fita única, DNA girase, proteína de iniciação.

Alongamento

Na fase de alongamento da replicação, o DNA de fita única é usado como um molde para a síntese de DNA. Esse processo requer várias enzimas.

Síntese dos *primers*. Todas as DNA polimerases precisam de um nucleotídio com um grupo 3′-OH ao qual um novo nucleotídio pode ser adicionado. Por causa dessa exigência, as DNA polimerases não podem iniciar a síntese de DNA em um molde vazio; elas precisam de um *primer* – um grupo 3′-OH – para iniciar. Então, como começa a síntese de DNA?

Uma enzima chamada **primase** sintetiza segmentos curtos (cerca de 10 a 12 nucleotídios) de nucleotídios de RNA, ou ***primers***, que fornecem um grupo 3′-OH ao qual a DNA polimerase pode prender os nucleotídios de DNA. (Como a primase é uma RNA polimerase, ela não requer um grupo 3′-OH existente para iniciar a síntese de uma fita de nucleotídios.) Todas as moléculas de DNA inicialmente têm *primers* de RNA curtos incrustados dentro delas; esses *primers* são removidos e substituídos por nucleotídios de DNA.

Na fita líder, onde a síntese de DNA é contínua, um *primer* é necessário apenas na extremidade 5′ da fita recém-sintetizada. Na fita tardia, onde a replicação é descontínua, um novo *primer* deve ser gerado no início de cada fragmento de Okazaki (**Figura 12.13**). A primase forma um complexo com a helicase na forquilha de replicação e se desloca ao longo do molde da fita tardia. É provável que o *primer* único na fita líder seja sintetizado pelo complexo primase-helicase no molde da fita tardia na *outra* forquilha de replicação, na extremidade oposta da bolha de replicação. Resolva o Problema 30

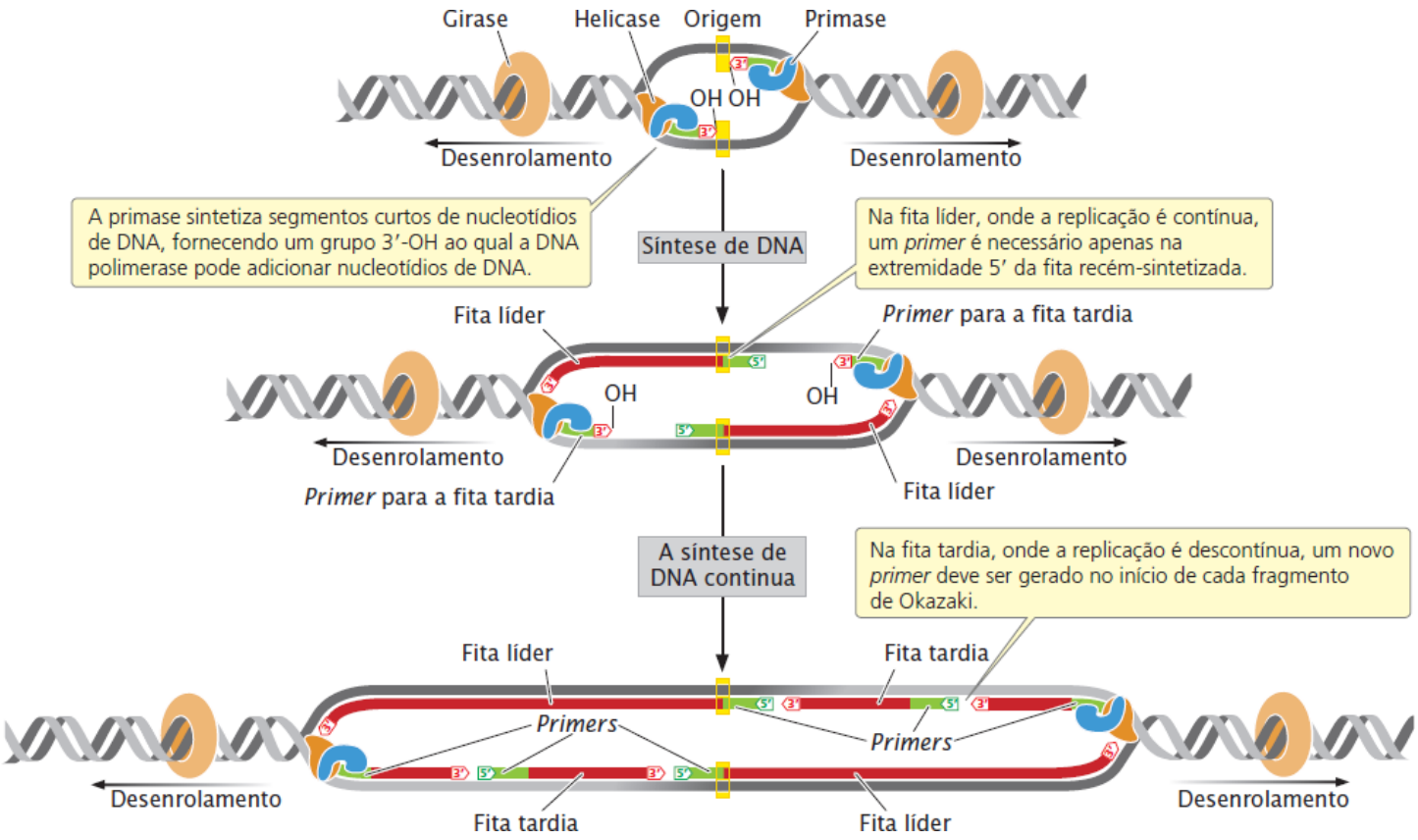


Figura 12.13 A primase sintetiza segmentos curtos de nucleotídios de DNA, fornecendo um grupo 3′-OH ao qual a DNA polimerase pode adicionar nucleotídios de DNA.

Conceitos

A primase sintetiza um segmento curto de nucleotídios de RNA (*primers*), que fornece um grupo 3′-OH para a fixação dos nucleotídios de DNA para início da síntese de DNA.

Checagem dos conceitos 5

Os *primers* são sintetizados em que região na fita tardia?

a. Apenas na extremidade 5’ da fita recém-sintetizada.
b. Apenas na extremidade 3’ da fita recém-sintetizada.
c. No início de cada fragmento de Okazaki.
d. Em múltiplos locais dento do fragmento de Okazaki.

Síntese de DNA pelas DNA polimerases. Após o DNA ser desenrolado e um *primer* ser adicionado, as DNA polimerases prolongam a fita nova de polinucleotídios ao catalisar a polimerização do DNA. As polimerases mais estudadas são da *E. coli*, que tem pelo menos cinco DNA polimerases diferentes. Duas delas, a DNA polimerase I e a DNA polimerase III, sintetizam DNA na replicação (**Quadro 12.3**); as outras três têm funções especializadas no reparo do DNA.

A **DNA polimerase III** é um grande complexo de multiproteínas que atua como ferramenta principal da replicação. Ela sintetiza fitas de nucleotídios ao adicionar novos nucleotídios à extremidade 3′ da molécula de DNA crescente. A DNA polimerase III tem duas atividades enzimáticas (ver Quadro 12.3). Sua atividade polimerase 5′ → 3′ possibilita que ela adicione novos nucleotídios no sentido 5′ → 3′. Sua atividade exonuclease 3′ → 5′ possibilita que ela remova nucleotídios na direção 3′ → 5′, tornando possível a correção de erros. Se um nucleotídio com uma base incorreta é inserido na molécula de DNA crescente, a DNA polimerase III usa a atividade exonuclease 3′ → 5′ para apoiar e remover o nucleotídio incorreto. Ela então reinicia sua atividade polimerase 5′ → 3′. Essas duas funções juntas fazem a DNA polimerase III sintetizar de modo eficiente e preciso novas moléculas de DNA. A DNA polimerase III apresenta alta processividade, sendo, portanto, capaz de adicionar vários nucleotídios à fita de DNA crescente sem liberar o molde: normalmente ela segura o molde e continua a sintetizar o DNA até que o molde tenha sido completamente replicado. A alta processividade da DNA polimerase III é garantida por um dos polipeptídios que compõem a enzima. Esse polipeptídio, chamado de subunidade β, serve como um grampo para a enzima: ele circula o DNA e mantém a DNA polimerase ligada à fita molde durante a replicação. A DNA polimerase III adiciona nucleotídios de DNA ao *primer*, sintetizando o DNA das fitas líder e tardia.

A primeira polimerase da *E. coli* a ser descoberta, a **DNA polimerase I**, também apresenta atividades polimerase 5′ → 3′ e exonuclease 3′ → 5′ (ver Quadro 12.3), tornando possível que a enzima sintetize DNA e corrija erros. Diferente da DNA polimerase III, esta polimerase também tem atividade exonuclease 5′ → 3′, que é usada para remover os *primers* entregues pela primase e então os substitui por nucleotídios de DNA ao sintetizar na direção 5′ → 3′. A DNA polimerase I tem processividade menor que a DNA polimerase III. A remoção e substituição dos *primers* parece ser a função principal da DNA polimerase I. Após a DNA polimerase III iniciar a síntese no *primer* e mover-se na direção 3′ (*downstream*), a DNA polimerase I remove os nucleotídios de RNA do *primer*, substituindo-os por nucleotídios de DNA. As DNA polimerases II, IV e V funcionam no reparo de DNA.

Apesar de suas diferenças, todas as polimerases da *E. coli:*

1. Sintetizam qualquer sequência especificada pela fita molde.
2. Sintetizam na direção 5′ → 3′ ao adicionar nucleotídios a um grupo 3′-OH.
3. Usam dNTPs para sintetizar DNA novo.
4. Precisam de um grupo 3′-OH para iniciar a síntese.
5. Catalisam a formação de uma ligação fosfodiéster ao unir o grupo 5′-fosfato do nucleotídio que chega ao grupo 3′-OH do nucleotídio anterior na fita crescente, rompendo dois fosfatos no processo.
6. Produzem fitas recém-sintetizadas que são complementares e antiparalelas às fitas molde.
7. Estão associadas a várias outras proteínas. Resolva o Problema 27

Conceitos

As DNA polimerases sintetizam DNA no sentido 5′ →3′ ao adicionar novos nucleotídios à extremidade 3′ de uma fita de nucleotídios crescente.

DNA ligase. Após a DNA polimerase III se ligar ao nucleotídio do grupo 3′-OH no último nucleotídio do *primer* de RNA, cada novo nucleotídio de DNA fornece o grupo 3′-OH necessário para que o próximo nucleotídio seja adicionado. Esse processo continua enquanto o molde estiver disponível (**Figura 12.14 A**). A DNA polimerase I segue a DNA polimerase III e, ao usar sua atividade exonuclease 5′ → 3′, ela remove o *primer* de RNA. Ela então usa a atividade polimerase 5′ → 3′ para substituir os nucleotídios de RNA por nucleotídios de DNA. A DNA polimerase I se fixa ao primeiro nucleotídio do grupo OH

na extremidade 3′ do fragmento de Okazaki anterior e então continua no sentido 5′ → 3′ ao longo da fita de nucleotídios e substituindo, um por vez, os nucleotídios de DNA do *primer* (**Figura 12.14 B**).

Quadro 12.3 Características das DNA polimerases na *E. coli*.

DNA polimerase	Polimerização 5′ → 3′	Exonuclease 3′ → 5′	Exonuclease 5′ → 3′	Função
I	Sim	Sim	Sim	Remove e substitui os *primers*
II	Sim	Sim	Não	Repara o DNA; reinicia a replicação depois de o DNA danificado parar a síntese
III	Sim	Sim	Não	Alonga o DNA
IV	Sim	Não	Não	Repara o DNA
V	Sim	Não	Não	Repara o DNA; sintetiza o DNA translesão

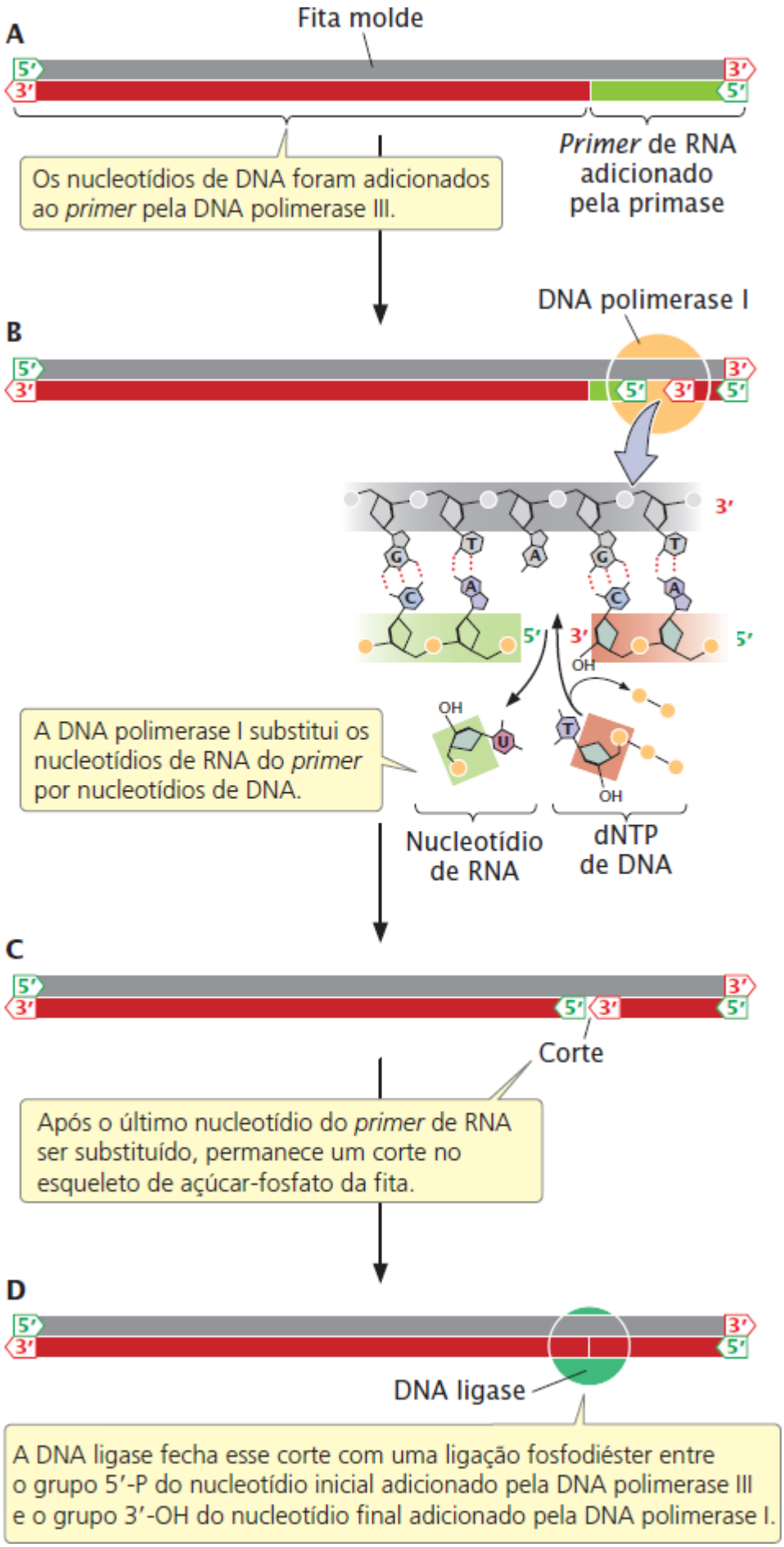


Figura 12.14 A DNA ligase fecha a ruptura deixada pela DNA polimerase I na estrutura açúcar-fosfato.

Após a polimerase I ter substituído o último nucleotídio do *primer* de RNA por um nucleotídio DNA, permanece uma ruptura na estrutura açúcar-fosfato da nova fita de DNA. O grupo 3′-OH do último nucleotídio a ser adicionado pela DNA polimerase I não está preso ao grupo 5′-fosfato do primeiro nucleotídio adicionado pela DNA polimerase III (**Figura 12.14 C**). Essa ruptura

é fechada pela enzima **DNA ligase**, que catalisa a formação de uma ligação fosfodiéster sem adicionar outro nucleotídio à fita (**Figura 12.14 D**). Algumas das principais enzimas e proteínas necessárias para a replicação do DNA procariótico estão resumidas no **Quadro 12.4**.

Conceitos

Após os *primers* serem removidos e substituídos, a ruptura na ligação açúcar-fosfato é fechada pela DNA ligase.

Checagem dos conceitos 6

Qual enzima bacteriana remove os *primers*?

a. Primase.
b. DNA polimerase I.
c. DNA polimerase III.
d. Ligase.

Alongamento na forquilha de replicação. Agora que os principais componentes enzimáticos do alongamento – DNA polimerases, helicase, primase e ligase – foram apresentados, analisemos como eles interagem na forquilha de replicação. Como a síntese de ambas as fitas ocorre simultaneamente, tem de haver duas unidades da DNA polimerase III na forquilha de replicação, uma para cada fita. Em um modelo do processo de replicação, as duas unidades de DNA polimerase III estão conectadas (**Figura 12.15**), o molde da fita tardia se fecha ao redor de modo que fica na posição para replicação 5′ → 3′. Dessa maneira, o complexo da DNA polimerase III é capaz de fazer a replicação 5′ → 3′ simultaneamente em ambos os moldes, mesmo que eles estejam em direções opostas. Após cerca de 1.000 pb de DNA novo terem sido sintetizados, a DNA polimerase III libera o molde da fita tardia e uma nova alça se forma (ver Figura 12.15). A primase sintetiza um novo *primer* na fita tardia, e a DNA polimerase III, então, sintetiza um novo fragmento de Okazaki.

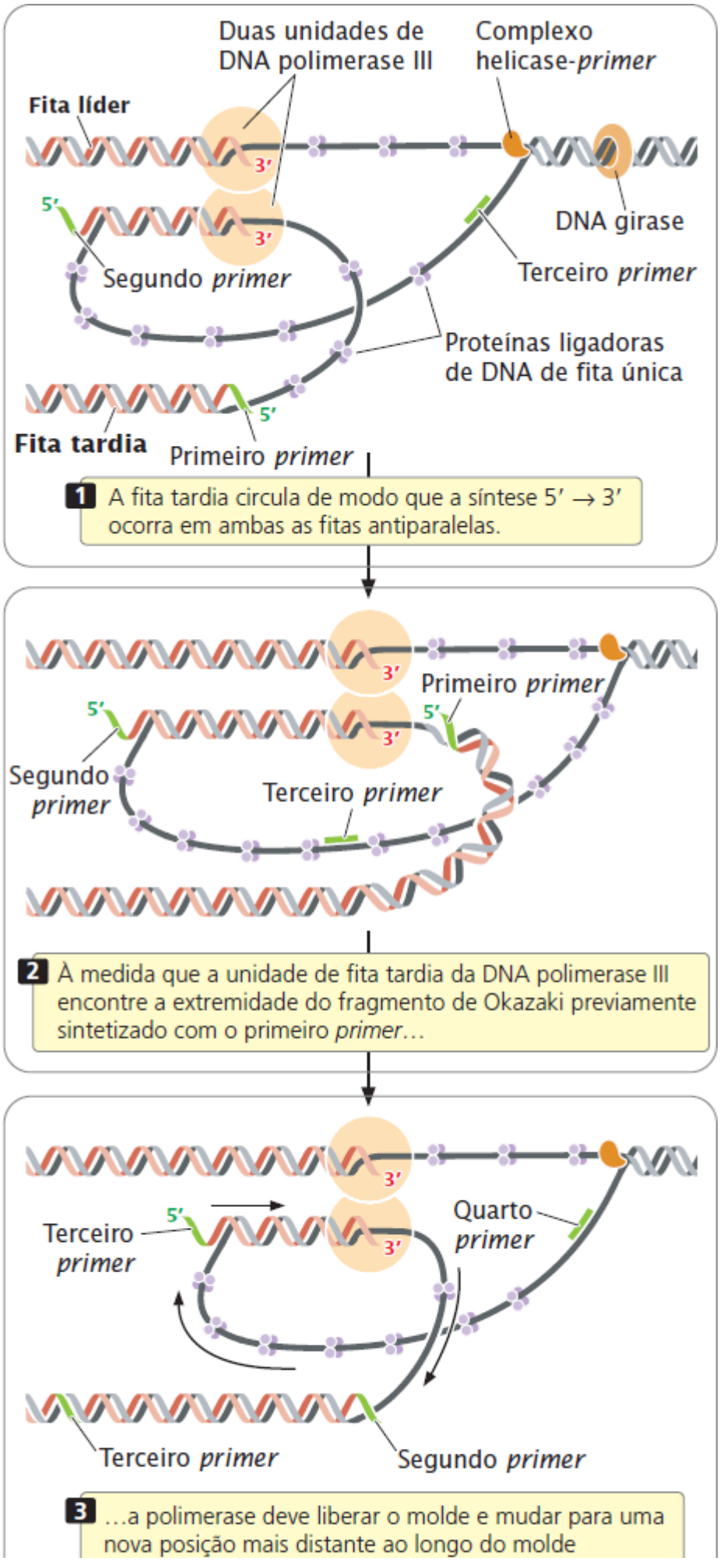
Quadro 12.4 Componentes necessários para a replicação nas células bacterianas.

Componente	Função
Proteína de iniciação	Liga-se à origem e separa as fitas do DNA para iniciar a replicação
DNA helicase	Desenrola o DNA na forquilha de replicação
Proteínas ligadoras de DNA de fita única	Prendem-se ao DNA de fita única e evitam a formação de estruturas secundárias
DNA girase	Desloca-se na forquilha de replicação, abrindo e fechando rupturas no DNA de dupla-hélice para liberar o torque resultante do desenrolamento na forquilha de replicação
DNA primase	Sintetiza um *primer* de RNA curto para fornecer um grupo 3′-OH para a fixação dos nucleotídios de DNA
DNA polimerase III	Alonga a fita nova de nucleotídios a partir do grupo 3′-OH fornecido pelo *primer*
DNA polimerase I	Remove os *primers* de RNA e os substitui por DNA
DNA ligase	Une os fragmentos de Okazaki ao fechar as rupturas no esqueleto de açúcar-fosfato do DNA recém-sintetizado

Em resumo, cada forquilha de replicação ativa requer cinco componentes básicos:

1. A helicase para desenrolar o DNA.

2. As proteínas ligadoras de fita única para proteger as fitas de nucleotídios e evitar a formação de estruturas secundárias.



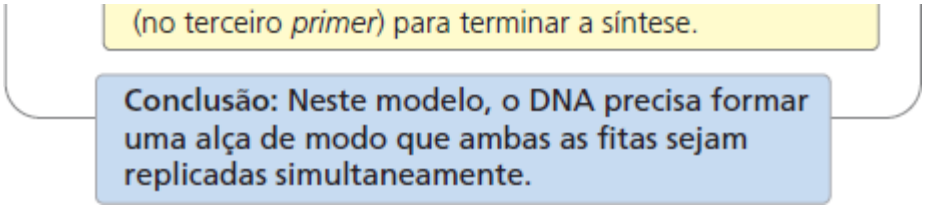


Figura 12.15 Em um modelo de replicação de DNA na *E. coli*, as duas unidades de DNA polimerase III estão conectadas. O molde de fita tardia forma uma alça de modo que a replicação ocorra nas duas fitas de DNA antiparalelas. Os componentes do maquinário de replicação na forquilha de replicação estão apresentados na parte superior.

3. A topoisomerase girase para remover a fita à frente da forquilha de replicação.
4. A primase para sintetizar *primers* com um grupo 3′-OH no início de cada fragmento de DNA.
5. A DNA polimerase para sintetizar a fita líder e fita tardia de nucleotídios.

Término

Em algumas moléculas de DNA, a replicação é terminada sempre que duas forquilhas de replicação se encontram. Em outras, as sequências de término específicas (chamadas sítios *Ter*) bloqueiam a replicação. Uma proteína de término, chamada Tus na *E. coli*, liga-se a essas sequências, criando um complexo Tus-*Ter* que bloqueia o movimento da helicase, paralisando a forquilha de replicação e evitando a replicação de DNA. Cada complexo Tus-*Ter* bloqueia a forquilha de replicação de se deslocar em um sentido, mas não para o outro.

Fidelidade da replicação do DNA

Em geral, a taxa de erro na replicação é menor que um erro por bilhão de nucleotídios. Como é alcançada essa incrível exatidão?

As DNA polimerases são muito peculiares no pareamento dos nucleotídios com seus complementos na fita molde. Surgem erros na seleção de nucleotídios pela DNA polimerase em uma proporção de apenas uma vez a cada 100.000 nucleotídios. A maioria dos erros que surgem na seleção dos nucleotídios é corrigida em um segundo processo, chamado **revisão**. Quando uma DNA polimerase insere um nucleotídio incorreto na fita crescente, o grupo 3′-OH do nucleotídio mal pareado não está corretamente posicionado no sítio ativo da DNA polimerase para aceitar o próximo nucleotídio. O posicionamento incorreto atrasa a reação de polimerização e a atividade exonuclease 3′ → 5′ da DNA polimerase remove o nucleotídio incorretamente pareado. A DNA polimerase, então, insere o nucleotídio correto. Juntos, a revisão e a seleção de nucleotídios resultam em uma taxa de erro de apenas um em 10 milhões de nucleotídios.

Um terceiro processo, chamado de **reparo de pareamento errado** de DNA (discutido com mais detalhes no *Capítulo 18*), corrige os erros após o término da replicação. Qualquer nucleotídio incorretamente pareado que permanece após a replicação produz uma deformação na estrutura secundária do DNA; a deformidade é identificada por enzimas que removem o nucleotídio incorreto e usam a fita original de nucleotídios como molde para substituir esse nucleotídio. O reparo do pareamento errado requer a capacidade de distinguir entre as fitas antiga e nova de DNA, porque as enzimas precisam de uma maneira para determinar qual das duas bases incorretamente pareadas deve ser removida. Na *E. coli*, grupos metila ($–CH_3$) são adicionados às sequências específicas de nucleotídios, mas apenas *depois* da replicação. Assim, imediatamente após a síntese do DNA, apenas a fita de DNA antiga é metilada. Ela pode ser diferenciada a partir da fita recém-sintetizada, e o reparo do pareamento errado ocorre preferencialmente na fita de nucleotídios não metilada. Nenhum processo isolado poderia produzir esse grau de exatidão; são necessários vários processos, e cada um deles identifica os erros não detectados pelos processos anteriores.

Conceitos

A replicação é um processo muito exato, com menos de um erro por um bilhão de nucleotídios. O nível elevado de exatidão na replicação do DNA é obtido pela seleção dos nucleotídios, revisão e reparo de pareamento errado.

Checagem dos conceitos 7

Qual mecanismo consegue diferenciar a fita recém-sintetizada da fita-molde do DNA?

a. Seleção de nucleotídios.
b. Revisão do DNA.
c. Reparo de pareamento errado.
d. Todas as opções anteriores.

Conceitos conectantes

As regras básicas da replicação

A replicação bacteriana requer várias enzimas (ver Quadro 12.4), proteínas e sequências de DNA que trabalham juntas para sintetizar uma nova molécula de DNA. Esses componentes são importantes, mas não nos aprofundamos nos detalhes do processo para não nos perdermos dos princípios gerais da replicação.

1. A replicação é sempre semiconservativa.
2. A replicação sempre inicia em sequências chamadas de origens.
3. A síntese do DNA é iniciada por segmentos curtos de RNA chamados de *primers*.
4. O alongamento das fitas de DNA é sempre na direção 5′ → 3′.
5. O DNA novo é sintetizado a partir de dNTPs; na polimerização do DNA, dois grupos de fosfato são rompidos a partir de um dNTP e o nucleotídio resultante é adicionado ao grupo 3′-OH da fita crescente de nucleotídios.
6. A replicação é contínua na fita líder e descontínua na fita tardia.
7. As fitas novas de nucleotídios são complementares e antiparalelas em relação a suas fitas molde.
8. A replicação ocorre em taxas muito elevadas e é muito precisa, graças à exata seleção de nucleotídios, revisão e reparo do pareamento errado.

12.4 A replicação do DNA eucariótico é semelhante à replicação bacteriana, mas existem algumas diferenças

Embora a replicação eucariótica lembre a replicação bacteriana em muitos aspectos, a replicação nas células eucarióticas apresenta muitos desafios adicionais. Primeiro, o tamanho muito maior dos genomas eucarióticos exige que a replicação seja iniciada em múltiplas origens. Segundo, os cromossomos eucarióticos são lineares, enquanto os procarióticos são circulares. Terceiro, o molde de DNA está associado a proteínas histonas na forma de nucleossomos, e a montagem dos nucleossomos deve ser feita imediatamente à replicação do DNA.

Origens eucarióticas

Os pesquisadores isolaram pela primeira vez as origens eucarióticas da replicação em células de levedura ao demonstrar que algumas sequências de DNA possibilitam a replicação quando transferidas de um cromossomo de levedura para pequenos pedaços circulares de DNA (plasmídios). Essas **sequências de replicação autônoma** (ARSs) possibilitam a replicação de qualquer DNA ao qual estejam ligadas. Posteriormente, foi demonstrado que elas são as origens da replicação nos cromossomos de levedura. As origens da replicação de diferentes organismos eucarióticos variam muito na sequência, embora elas tenham vários pares de base A-T. Um complexo de multiproteínas, o complexo de reconhecimento da origem (ORC), liga-se às origens e desenrola o DNA nessa região.

Conceitos

O DNA eucariótico tem muitas origens de replicação. Em cada origem, um complexo multiproteico de reconhecimento da origem se liga para iniciar o desenrolamento do DNA.

Checagem dos conceitos 8

Em comparação com os procariotos, quais são as diferenças na estrutura do genoma das células eucarióticas que afetam a maneira como a replicação ocorre?

Licenciamento da replicação do DNA

As células eucarióticas utilizam milhares de origens, de modo que o genoma inteiro possa ser replicado em tempo hábil. Entretanto, o uso de múltiplas origens cria um problema especial no cronograma da replicação: o genoma inteiro precisa ser replicado com precisão uma vez, e apenas uma vez, em cada ciclo celular de modo que nenhum gene seja esquecido e nenhum gene seja replicado mais de uma vez. Como uma célula garante que a replicação seja iniciada em milhares de origens apenas uma vez por ciclo celular?

A replicação precisa do DNA é possível ao separar a iniciação da replicação em duas etapas diferentes. Na primeira etapa, as

origens são licenciadas, aprovadas para replicação. Isso ocorre no início do ciclo celular, quando um **fator de licenciamento de replicação** se prende a uma origem. Na segunda etapa, o maquinário da replicação inicia o processo em cada origem *licenciada*. A chave é que o maquinário de replicação atue apenas nas origens licenciadas. À medida que as forquilhas de replicação se afastam da origem, o fator de licenciamento é removido, deixando a origem em um estado de não licenciada, em que a replicação não pode ser iniciada novamente até a renovação do licenciamento. Para garantir que a replicação ocorra apenas uma vez por ciclo celular, o fator de licenciamento é ativado apenas depois de a célula ter completado a mitose e antes que a replicação seja iniciada.

Um fator de licenciamento eucariótico é um complexo chamado MCM (para manutenção de minicromossomo), com uma DNA helicase que desenrola um segmento curto de DNA na iniciação da replicação. O MCM deve se ligar ao DNA para iniciar a replicação em uma origem. Após a replicação começar em uma origem, uma proteína chamada Geminina evita que o MCM se ligue ao DNA e reinicie a replicação naquela origem. No final da mitose, a Geminina é degradada, tornando possível que o MCM se ligue mais uma vez ao DNA e autorize novamente a origem. O MCM também funciona como DNA helicase durante o processo de replicação.

Desenrolamento

Foram isoladas várias helicases diferentes que separam o DNA de fita dupla nas células eucarióticas, como as proteínas ligadoras de fita única e as topoisomerases (que têm uma função equivalente à DNA girase nas células bacterianas). Essas enzimas e proteínas atuam no desenrolamento do DNA eucariótico da mesma maneira que as correspondentes bacterianas.

DNA polimerases eucarióticas

Algumas diferenças expressivas nos processos de replicação bacteriana e eucariótica são o número e as funções das DNA polimerases. As células eucarióticas têm várias DNA polimerases diferentes que atuam na replicação, na recombinação e no reparo de DNA.

Três DNA polimerases fazem a maior parte da síntese do DNA durante a replicação: DNA polimerase α, DNA polimerase δ e DNA polimerase ε (**Quadro 12.5**). A **DNA polimerase α** tem atividade primase e inicia a síntese de DNA nuclear ao sintetizar um *primer* de RNA, seguida por uma sequência curta de nucleotídios de DNA. Após a DNA polimerase α ter se estabelecido entre 30 e 40 nucleotídios, a **DNA polimerase δ** completa a replicação na fita tardia. Semelhante em estrutura e função à DNA polimerase δ, a **DNA polimerase ε** replica a fita líder. Outras DNA polimerases participam da replicação e recombinação ou catalisam a replicação do DNA de organela.

Algumas DNA polimerases, como a DNA polimerase δ e a DNA polimerase ε, são capazes de replicar o DNA em alta velocidade e com grande fidelidade (poucos erros) porque têm sítios ativos que acomodam de modo justo e exclusivo os quatro nucleotídios normais do DNA, os monofosfatos de adenosina, guanosina, citidina e timidina. Como resultado dessa especificidade, moldes destorcidos de DNA e bases anormais não são acomodados com facilidade dentro do sítio ativo da enzima. Quando esses erros são encontrados no molde de DNA, as DNA polimerases de alta fidelidade param e não conseguem contornar a lesão.

Outras DNA polimerases têm menor fidelidade, mas são capazes de contornar as distorções no molde de DNA. Essas **DNA polimerases translesão** especializadas têm um sítio ativo mais aberto e são capazes de acomodar e copiar moldes com bases anormais, estruturas distorcidas e lesões volumosas. Assim, essas enzimas especializadas podem contornar tais erros, mas como seus sítios ativos são mais abertos e acomodados, elas tendem a produzir mais erros. Na replicação, as enzimas de alta velocidade e alta fidelidade são usadas até encontrarem um bloqueio da replicação. Nesse ponto, uma ou mais polimerases translesão assumem, contornam a lesão e continuam replicando uma seção curta de DNA, Então, as polimerases translesão se separam da forquilha de replicação e as enzimas de alta fidelidade encerram a replicação com alta velocidade e exatidão. As enzimas de reparo de DNA reparam os erros produzidos pelas polimerases translesão, embora alguns desses erros possam escapar da detecção e levem a mutações.

Conceitos

Existem muitas DNA polimerases diferentes nas células eucarióticas. As DNA polimerases α, δ e ε fazem a replicação na fita líder e na fita tardia. Outras DNA polimerases reparam o DNA. As polimerases translesão especializadas são usadas para contornar as distorções do molde de DNA que normalmente paralisam as principais DNA polimerases.

Checagem dos conceitos 9

Algumas das DNA polimerases eucarióticas tendem a fazer erros na replicação. Por que uma célula usaria uma DNA polimerase suscetível a erro em vez de uma mais exata?

Montagem do nucleossomo

O DNA eucariótico é complexado a proteínas histonas em estruturas chamadas nucleossomos, que contribuem para a estabilidade e o empacotamento da molécula de DNA (ver Figura 11.3). Na replicação, a estrutura da cromatina é rompida pela forquilha de replicação, mas os nucleossomos são rapidamente remontados em duas novas moléculas de DNA. As micrografias eletrônicas do DNA eucariótico, como a apresentada na **Figura 12.16**, mostram DNA recém-replicado já coberto por nucleossomos, indicando que os nucleossomos são remontados rapidamente.

Quadro 12.5 DNA polimerases nas células eucarióticas.

DNA polimerase	Atividade polimerase 5′ → 3′	Atividade exonuclease 3′ → 5′	Função celular
α (alfa)	Sim	Não	Iniciação da síntese de DNA nuclear e reparo do DNA; tem atividade da primase
δ (delta)	Sim	Sim	Síntese de fita tardia do DNA nuclear, reparo do DNA e síntese do DNA translesão
ε (épsilon)	Sim	Sim	Síntese da fita líder
γ (gama)	Sim	Sim	Replicação e reparo do DNA mitocondrial
ξ (zeta)	Sim	Não	Síntese de DNA translesão
η (eta)	Sim	Não	Síntese de DNA translesão
θ (teta)	Sim	Não	Reparo do DNA
ι (iota)	Sim	Não	Síntese de DNA translesão
κ (capa)	Sim	Não	Síntese de DNA translesão
λ (lambda)	Sim	Não	Reparo do DNA
μ (mi)	Sim	Não	Reparo do DNA
σ (sigma)	Sim	Não	Replicação do DNA nuclear (possivelmente), reparo do DNA e coesão da cromátide-irmã
ϕ (fi)	Sim	Não	Síntese de DNA translesão
Rev1	Sim	Não	Reparo do DNA

Nota: As polimerases listadas na parte superior do quadro são as enzimas responsáveis pela replicação do DNA.

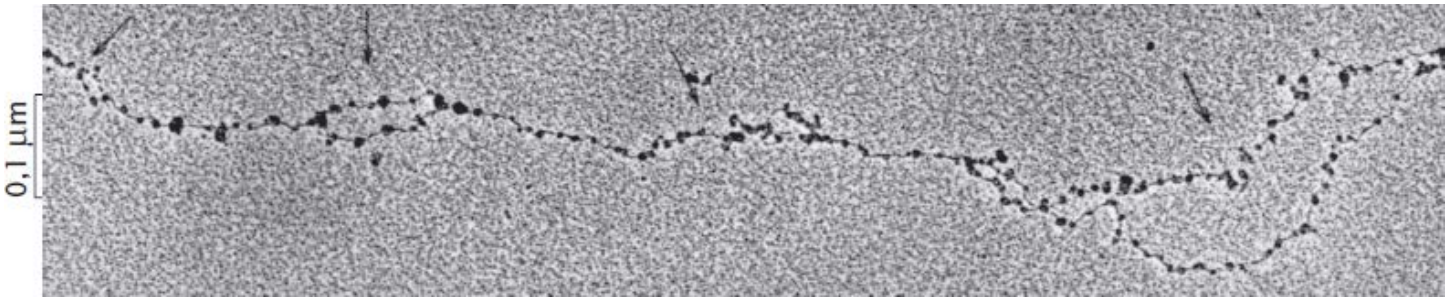


Figura 12.16 Os nucleossomos são rapidamente remontados no DNA recém-sintetizado. Esta micrografia eletrônica do DNA eucariótico no processo de replicação mostra claramente que o DNA recém-replicado já está coberto por nucleossomos (círculos escuros). (Victoria Foe.)

São necessárias três etapas para a criação de novos nucleossomos: (1) o rompimento dos nucleossomos originais na molécula de DNA parental à frente da forquilha de replicação; (2) a redistribuição das histonas preexistentes sobre novas moléculas de DNA e (3) a adição de histonas recém-sintetizadas para completar a formação de novos nucleossomos. Antes da replicação, uma molécula de DNA única está associada a proteínas histonas. Após a replicação e a montagem do nucleossomo, duas moléculas de DNA estão associadas às histonas. As histonas originais de um nucleossomo permanecem juntas, fixas a uma das novas moléculas de DNA, ou elas se desmontam e se misturam com histonas novas em ambas as moléculas de DNA?

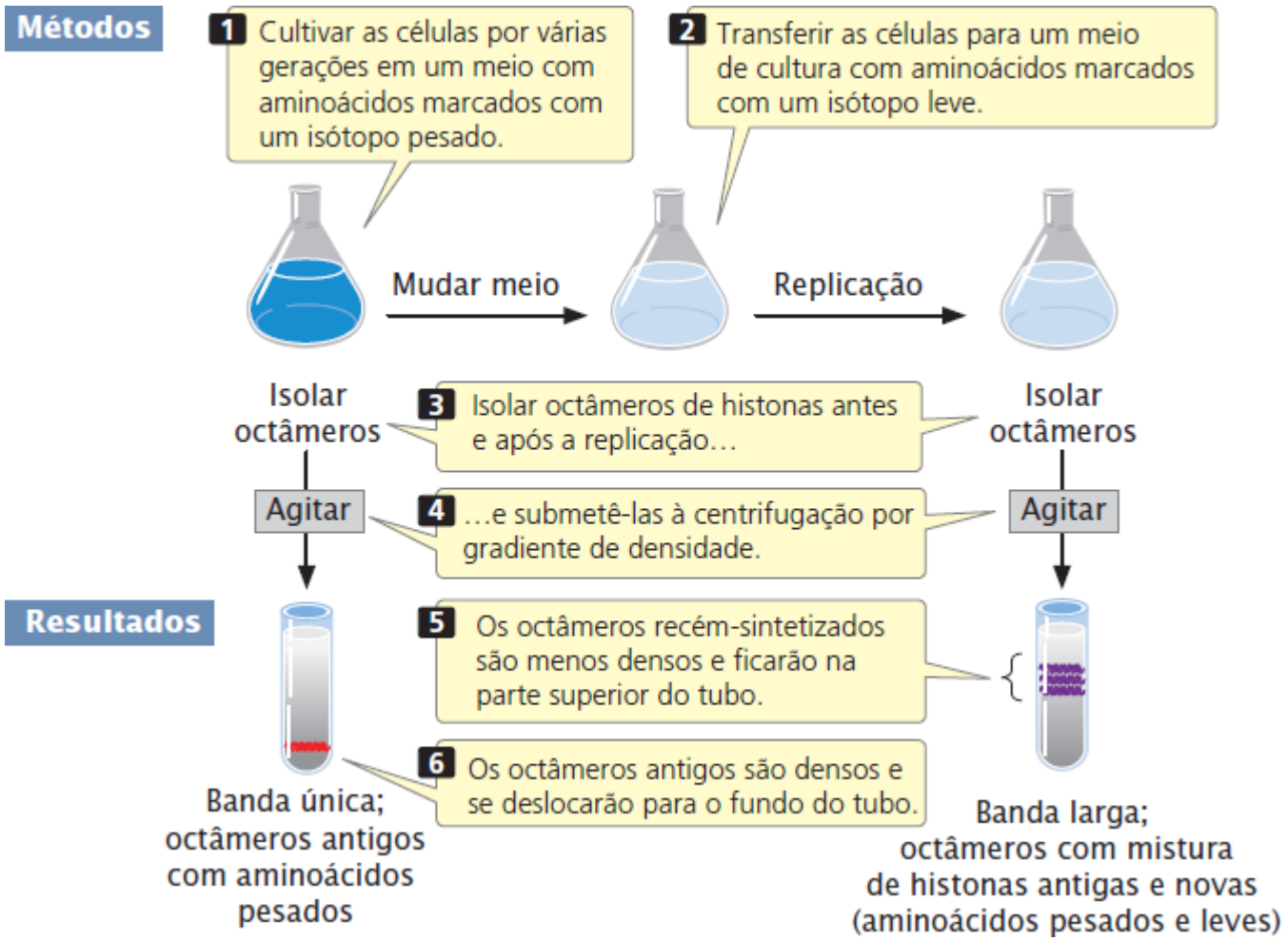
Técnicas semelhantes às empregadas por Meselson e Stahl para determinar o modo de replicação do DNA foram usadas para estudar este assunto. As células foram cultivadas por várias gerações em um meio com aminoácidos marcados com um isótopo pesado. As proteínas histonas incorporaram esses aminoácidos pesados e ficaram densas (**Figura 12.17**). As células foram, então, transferidas para um meio de cultura com aminoácidos marcados com um isótopo leve. As histonas montadas após a transferência tinham aminoácidos novos, leves, e eram menos densas.

Após a replicação, os octâmeros de histonas foram isolados e centrifugados em um gradiente de densidade. Os resultados mostraram que, após a replicação, os octâmeros eram uma banda contínua entre a alta densidade (representando os octâmeros antigos) e baixa densidade (representado os octâmeros novos). Esse achado indica que as estruturas recém-montadas são uma mistura de histonas antigas e novas. Outras evidências indicam que os nucleossomos reconstituídos aparecem nas novas moléculas de DNA logo após o novo DNA surgir a partir do maquinário de replicação.

A remontagem dos nucleossomos durante a replicação é facilitada por proteínas denominadas chaperonas de histona, associadas à enzima helicase que desenrola o DNA. As chaperonas de histonas aceitam as histonas antigas oriundas da molécula de DNA original e as depositam, junto com as histonas recém-sintetizadas, nas duas novas moléculas de DNA. A evidência atual sugere que o nucleossomo original é degradado em dois dímeros H2A-H2B (cada dímero com um H2A e um H2B) e um tetrâmero H3-H4 único (cada tetrâmero com duas histonas H3 e duas histonas H4). O tetrâmero H3-H4 antigo é então aleatoriamente transferido para uma das novas moléculas de DNA e serve como a fundação na qual as cópias novas e antigas dos dímeros de H2A-H2B são adicionadas. Os tetrâmeros H3-H4 recém-sintetizados e os dímeros H2A e H2B também são adicionados a cada nova molécula de DNA para completar a formação de novos nucleossomos. A montagem de novos nucleossomos é facilitada por uma proteína chamada fator 1 de montagem de cromatina (CAF-1). **Resolva o Problema 33**

Experimento

Pergunta: O que acontece com as histonas na replicação do DNA eucariótico?



Conclusão: Após a replicação do DNA, os octâmeros recém-remontados são uma mistura aleatória de histonas antigas e novas.

Figura 12.17 Procedimento experimental para estudar como os nucleossomos se dissociam e se reassociam durante a replicação.

Conceitos

Após a replicação do DNA, novos nucleossomos são rapidamente remontados nas moléculas de DNA. Os nucleossomos são degradados durante a replicação e remontados a partir de uma mistura de histonas antigas e novas. A remontagem dos nucleossomos durante a replicação é facilitada pelas chaperonas de histonas e pelos fatores de montagem de cromatina.

Localização da replicação no núcleo

As DNA polimerases responsáveis pela replicação são descritas frequentemente como se deslocassem no molde de DNA da mesma maneira que um trem viaja sobre os trilhos. Evidências recentes sugerem outro conceito. Uma visão mais exata é que a polimerase está fixa em uma localização e o molde de DNA passa através dela, com as moléculas de DNA recém-sintetizadas surgindo na outra extremidade.

As técnicas de microscopia por fluorescência, capazes de revelar sítios ativos da síntese do DNA, evidenciam que a maior parte da replicação no núcleo de uma célula eucariótica ocorre em um número limitado de sítios fixos, denominados fábricas de replicação. As micrografias *time-lapse* revelam que o DNA recém-replicado é expulso desses sítios. Resultados semelhantes foram obtidos nas células bacterianas.

Síntese de DNA e ciclo celular

A replicação do DNA é contínua nas bactérias que se dividem rapidamente. Nas células eucarióticas, entretanto, a replicação é coordenada com o ciclo celular. A passagem pelo ciclo celular, incluindo o início da replicação, é controlada por pontos de verificação do ciclo celular. O importante ponto de verificação G_1/S (ver *Capítulo 2*) mantém o ciclo celular em G_1 até o DNA

estar pronto para ser replicado. Após passar por esse ponto de verificação, a célula entra na fase S e o DNA é replicado. O sistema de licenciamento da replicação garante, então, que o DNA não seja replicado novamente até a célula ter passado pela mitose.

Replicação nas extremidades dos cromossomos

Uma diferença fundamental entre as replicações eucariótica e bacteriana é que os cromossomos eucarióticos são lineares e têm extremidades. Como já foi afirmado, o grupo 3′-OH necessário para a replicação pelas DNA polimerases é fornecido no início da replicação por *primers* de RNA que são sintetizados pela primase. Essa solução é temporária porque, inevitavelmente, os *primers* têm de ser removidos e substituídos pelos nucleotídios de DNA. Em uma molécula de DNA circular, o alongamento do círculo fornece um grupo 3′-OH imediatamente à frente do *primer* (**Figura 12.18 A**). Após a remoção do *primer*, a substituição dos nucleotídios de DNA pode ser feita para esse grupo 3′-OH.

Problema da replicação da extremidade. Nos cromossomos lineares com múltiplas origens, o alongamento do DNA nos replicons adjacentes também fornece um grupo 3′-OH anterior a cada *primer* (**Figura 12.18 B**). Na extremidade de um cromossomo linear, entretanto, não existe trecho adjacente do DNA replicado para fornecer esse grupo 3′-OH importante. Quando o *primer* na extremidade do cromossomo for removido, ele não pode ser substituído pelos nucleotídios do DNA, criando uma lacuna na extremidade do cromossomo, sugerindo que o cromossomo deve ficar progressivamente menor em cada ciclo de replicação. O encurtamento do cromossomo significaria que, quando um organismo se reproduz, ele transmitiria os cromossomos mais curtos do que os que herdou. Os cromossomos ficariam mais curtos a cada nova geração e, consequentemente, perderiam a estabilidade. Essa questão é chamada de problema da replicação da extremidade. Na verdade, o encurtamento do cromossomo ocorre em algumas células somáticas, mas nos organismos unicelulares, células germinativas e células embrionárias iniciais eles não são encurtados nem se autodestroem. Então, como as extremidades dos cromossomos lineares são replicadas?

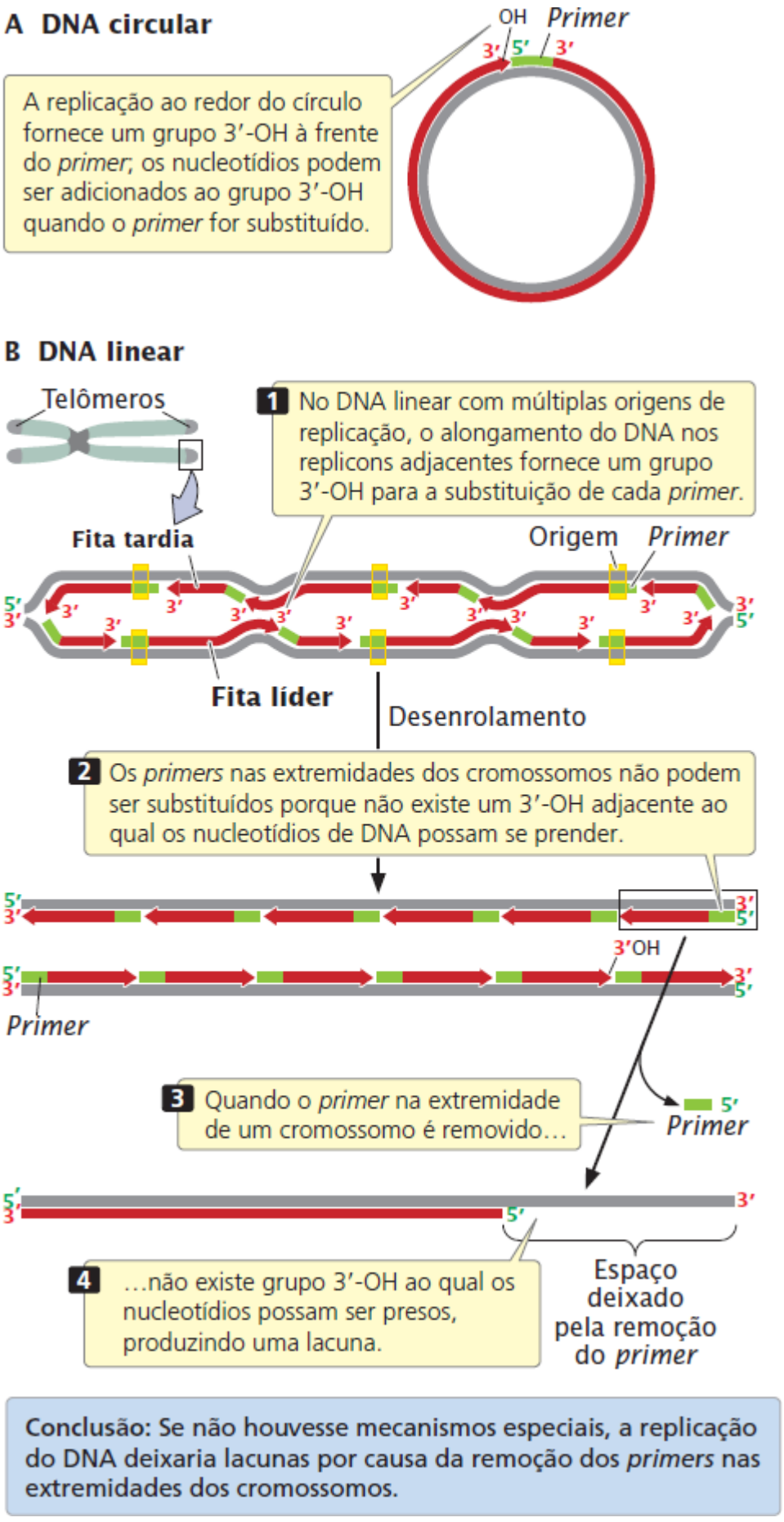


Figura 12.18 **A síntese de DNA nas extremidades dos cromossomos circulares e lineares é diferente.**

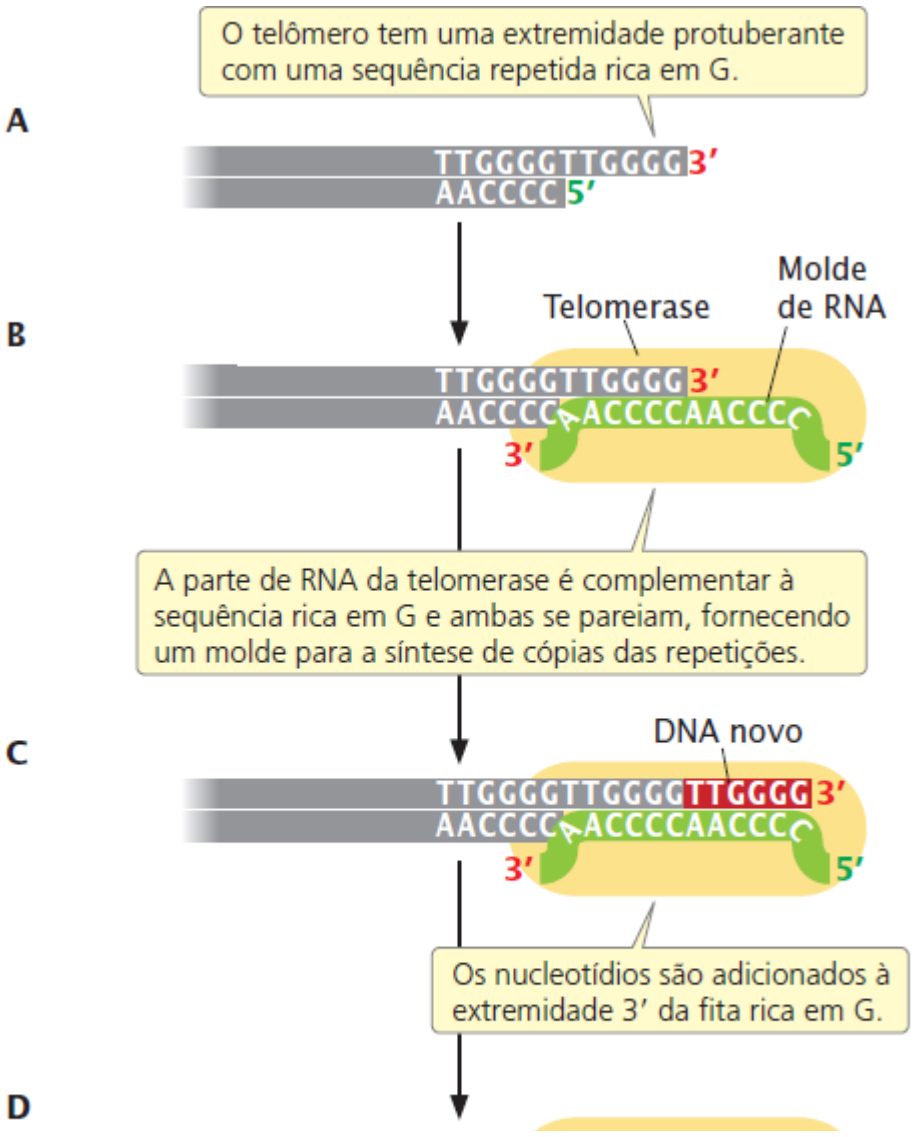
Telômeros e telomerase. As extremidades dos cromossomos – os telômeros – têm diversas características singulares, uma das quais é a variedade de cópias de uma sequência curta repetida. No protozoário *Tetrahymena* (em que essas sequências repetidas foram descobertas pela primeira vez) essa repetição telomérica é TTGGGG (ver Quadro 12.2), com essa fita rica em G se sobressaindo além da fita rica em C (**Figura 12.19 A**; veja também a seção sobre estrutura do telômero no *Capítulo 11*):

Direção do centrômero ← 5′-TTGGGGTTGGGG-3′ / 3′-AACCCC-5′ → Extremidade do cromossomo

A extremidade protrusa de fita única que se sobressai do telômero, conhecida como extremidade 3′ protuberante rica em G (***overhang* G**), pode ser estendida pela **telomerase**, uma enzima com um componente de proteína e um componente de RNA (também conhecida como uma ribonucleoproteína). A parte RNA da enzima tem de 15 a 22 nucleotídios que são complementares à sequência rica em G. Essa sequência pareia com a extremidade 3′ que sobressai do DNA (**Figura 12.19 B**) e fornece um molde para a síntese de cópias adicionais de DNA das repetições. Os nucleotídios de DNA são adicionados à extremidade 3′ da fita um por vez (**Figura 12.19 C**) e, após vários nucleotídios serem adicionados, o molde de RNA desloca o DNA e mais nucleotídios são adicionados à extremidade 3′ (**Figura 12.19 D**). Em geral, são adicionados de 14 a 16 nucleotídios na extremidade 3′ da fita rica em G.

Dessa forma, a telomerase pode expandir a extremidade 3′ do cromossomo sem o uso de um molde complementar de DNA (**Figura 12.19 E**). Ainda não está claro como a fita complementar rica em C é sintetizada (**Figura 12.19 F**). Sua síntese pode ocorrer por meio da replicação convencional, com a DNA polimerase α sintetizando um *primer* de RNA na extremidade 5′ do molde estendido (rico em G). A remoção desse *primer* novamente deixa um espaço na extremidade 5′ do cromossomo, mas esse espaço não é importante, porque a extremidade do cromossomo é estendida em cada replicação pela telomerase; então, o cromossomo não é encurtado.

A telomerase é encontrada em organismos unicelulares, células germinativas, células embrionárias iniciais e algumas células somáticas proliferativas (como as da medula óssea e as que revestem o intestino); todas têm obrigatoriamente divisão celular contínua. A maioria das células somáticas tem pouca ou nenhuma atividade de telomerase, e seus cromossomos são progressivamente encurtados em cada divisão celular. Essas células podem sofrer um número limitado de divisões; quando os telômeros são encurtados além de um ponto crítico, o cromossomo torna-se instável, com tendência a sofrer rearranjos, e é degradado. Esses eventos levam à morte celular.



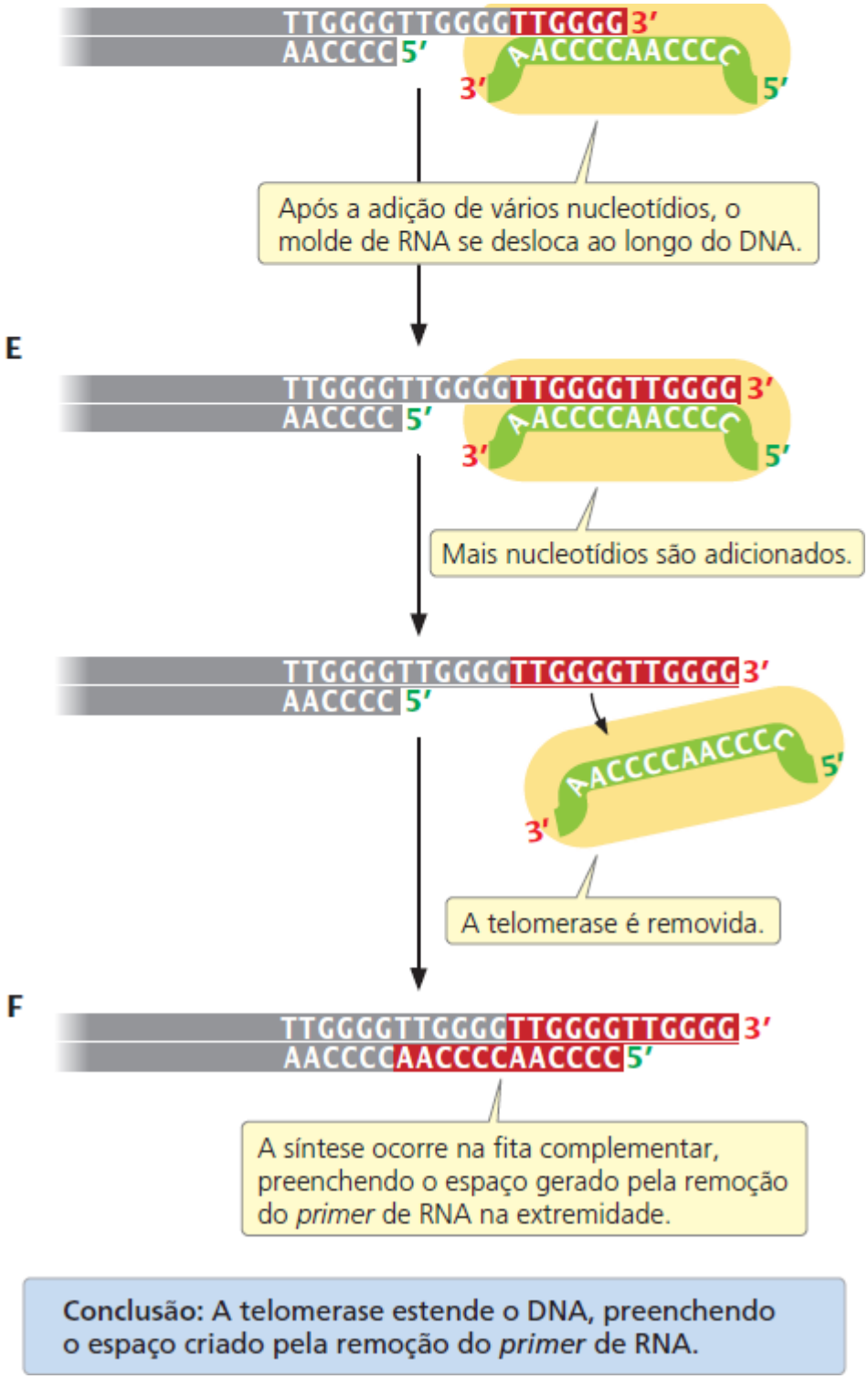


Figura 12.19 **A enzima telomerase é responsável pela replicação das extremidades do cromossomo.**

Conceitos

As extremidades dos cromossomos eucarióticos são replicadas por uma enzima RNA-proteína chamada telomerase. Essa enzima adiciona nucleotídios extras à fita de DNA rica em G do telômero.

Checagem dos conceitos 10

Qual seria o resultado se a telomerase de um organismo sofresse mutação e se tornasse não funcional?

a. Não ocorreria replicação do DNA.
b. A enzima DNA polimerase ficaria parada no telômero.
c. Os cromossomos seriam encurtados com cada nova geração.
d. Não seria possível remover os *primers* de RNA.

Telomerase, envelhecimento e doença. O encurtamento dos telômeros contribui para o processo de envelhecimento. Os telômeros de camundongos modificados pela engenharia genética sem um gene da telomerase funcional (e, portanto, que não expressam a telomerase nas células somáticas e germinativas) sofreriam encurtamento progressivo nas gerações sucessivas. Após várias gerações, esses camundongos apresentam alguns sinais de envelhecimento prematuro, como mudança da cor do pelo para grisalho, perda de pelo e cicatrização mais lenta de feridas. Por meio da engenharia genética também é possível criar células somáticas que expressam a telomerase. Nessas células, os telômeros não encurtam, o envelhecimento celular é inibido e as células se dividem indefinidamente.

Algumas das evidências mais fortes de que o comprimento do telômero está relacionado com o envelhecimento vieram de estudos dos telômeros em aves. Em 2012, cientistas do Reino Unido mediram o comprimento do telômero nas hemácias retiradas de 99 mandarins em vários períodos durante suas vidas. Os cientistas encontraram uma forte correlação entre o comprimento do telômero e a longevidade: os pássaros com telômeros mais longos viviam mais do que aqueles com telômeros curtos. A previsão mais forte de longevidade era quando o comprimento do telômero era medido no início da vida, aos 25 dias, que equivale aproximadamente à adolescência humana. Embora essas observações sugiram que o comprimento do telômero esteja associado ao envelhecimento em alguns animais, a precisão do seu papel no envelhecimento *humano* ainda é incerta.

Algumas doenças estão associadas a anormalidades da replicação do telômero. Pessoas portadoras da síndrome de Werner, uma doença autossômica recessiva, mostram sinais de envelhecimento prematuro que começam na adolescência ou no início da vida adulta, incluindo pele enrugada, cabelo grisalho, calvície, catarata e atrofia muscular. Elas desenvolvem câncer, osteoporose, cardiopatia e doença arterial e outras condições tipicamente associadas ao envelhecimento. O gene responsável, *WRN*, foi mapeado no cromossomo 8 humano e normalmente codifica uma enzima RecQ helicase. Essa enzima é necessária para a replicação eficiente dos telômeros. Nas pessoas com a síndrome de Werner, essa helicase é defeituosa e, consequentemente, os telômeros encurtam prematuramente.

Outra doença associada à manutenção anormal dos telômeros é a disqueratose congênita, que leva à insuficiência progressiva da medula óssea, caracterizada pela falha na produção de novas células sanguíneas pela medula. Pessoas com uma forma ligada ao X da doença apresentam uma mutação em um gene que codifica a discerina, uma proteína que normalmente ajuda a processar o componente RNA da telomerase. Essas pessoas herdam telômeros curtos de um genitor que carreia a mutação e que é incapaz de manter o comprimento do telômero em suas células germinativas em função de uma discerina defeituosa. Nas famílias que carreiam tal mutação, o comprimento do telômero encurta a cada geração, levando à antecipação – um aumento progressivo na intensidade da doença nas gerações seguintes (ver *Capítulo 5*).

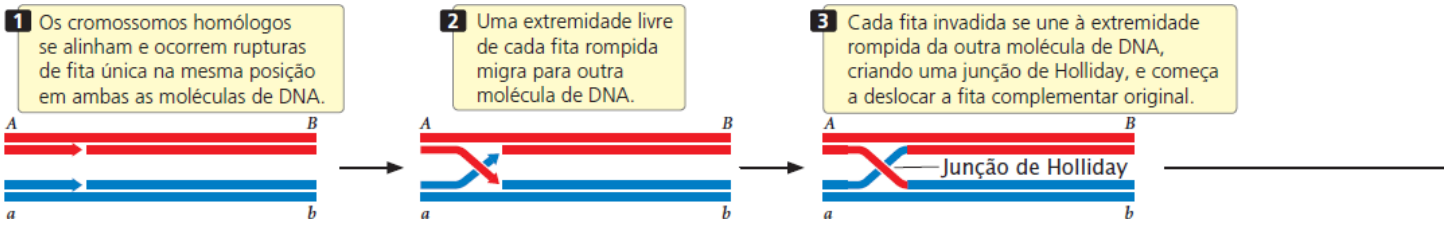
É possível que a telomerase participe no desenvolvimento do câncer. As células tumorais têm a capacidade de se dividir indefinidamente, e a telomerase é expressa em 90% de todos os cânceres. Algumas evidências recentes indicam que a telomerase estimula a proliferação celular independentemente do seu efeito no comprimento do telômero, mas o mecanismo usado pela telomerase para contribuir com o câncer não está claro. Como será discutido no Capítulo 23, o câncer é um processo complexo de várias etapas, que requer mutação em pelo menos vários genes. A ativação da telomerase sozinha não leva ao crescimento do câncer na maioria das células, mas parece que é necessário, junto com outras mutações, para surgir o câncer. Algumas drogas experimentais contra essa doença inibem a ação da telomerase.

Uma das dificuldades no estudo do efeito do encurtamento do telômero no processo de envelhecimento é que a expressão da telomerase nas células somáticas também promove câncer, o que pode encurtar o tempo de vida da pessoa. Para contornar esse problema, Antonia Tomas-Loba e seus colegas criaram camundongos modificados pela engenharia genética que expressam a telomerase e carreiam genes que os tornam resistentes ao câncer. Esses camundongos têm telômeros mais longos, vivem mais e exibem poucas mudanças relacionadas com a idade, como alterações na pele, redução da coordenação neuromuscular e doenças degenerativas. Esses resultados apoiam a ideia de que o encurtamento do telômero contribui para o envelhecimento.

Resolva o Problema 35

Replicação em archaea

O processo de replicação em archaea tem várias características em comum com a replicação nas células eucarióticas; muitas das proteínas envolvidas são mais semelhantes às proteínas encontradas nas células eucarióticas do que as das eubactérias. Como as eubactérias, algumas archaea têm uma única origem de replicação, mas a *Sulfolobus solfataricus* tem duas origens de replicação, semelhante às múltiplas origens observadas nos genomas eucarióticos. As origens da replicação em archaea não apresentam as sequências típicas identificadas pelas proteínas de iniciação bacterianas; pelo contrário, elas têm sequências semelhantes às encontradas nas origens eucarióticas. As proteínas de iniciação de archaea são mais parecidas com as proteínas dos eucariotos do que com as das eubactérias. Essas semelhanças na replicação entre as células de archaea e eucarióticas reforçam a conclusão de que as archaea estão mais próximas das células eucarióticas do que as eubactérias procarióticas.



12.5 Recombinação consiste em ruptura, alinhamento e reparo das fitas de DNA

A recombinação é a troca de informação genética entre as moléculas de DNA; quando a troca é entre moléculas de DNA homólogas, ela é chamada de **recombinação homóloga**. Esse processo ocorre no *crossing over*, no qual as regiões homólogas dos cromossomos são trocadas (ver Figura 7.5) e os alelos são misturados em novas combinações. A recombinação é um processo genético muito importante porque aumenta a variação genética. As taxas de recombinação fornecem informações importantes sobre as relações de ligação entre os genes, que são usadas para criar os mapas genéticos (ver Figuras 7.12 e 7.13). A recombinação também é essencial para alguns tipos de reparo de DNA (que serão discutidos no *Capítulo 18*).

A recombinação homóloga é um processo extraordinário: uma fita de nucleotídios de um cromossomo se alinha de maneira precisa com uma fita de nucleotídios do cromossomo homólogo; as rupturas surgem em regiões correspondentes de diferentes moléculas de DNA; as partes das moléculas trocam de lugar de modo preciso e, então, os fragmentos são reunidos corretamente. Nessa série complicada de eventos, não há perda ou ganho de informação genética. Embora o mecanismo molecular preciso da recombinação homóloga ainda não seja bem compreendido, é possível que a troca seja feita por meio do pareamento das bases complementares. Uma molécula de DNA de fita única de um cromossomo pareia com uma molécula de DNA de fita única de outro, formando o **DNA heteroduplex**.

Na meiose, a recombinação homóloga (*crossing over*) poderia ocorrer, teoricamente, antes, durante ou após a síntese de DNA. As evidências citológicas, bioquímicas e genéticas indicam que ela ocorre na prófase I da meiose, enquanto a replicação do DNA ocorre antes, na interfase. Assim, o *crossing over* exige a ruptura e a reunião das cromátides quando os cromossomos homólogos estão no estágio de quatro fitas (ver Figura 7.5). Esta seção explora algumas teorias de como o processo de recombinação ocorre.

Modelos de recombinação

A recombinação homóloga pode ocorrer por várias vias diferentes. Uma via é iniciada por uma ruptura de fita única em cada uma das duas moléculas de DNA e inclui a formação de uma estrutura especial chamada **junção de Holliday** (**Figura 12.20**). Nesse modelo, as moléculas de DNA de fita dupla de dois cromossomos homólogos se alinham de maneira precisa. Uma ruptura de fita única em uma das moléculas de DNA fornece uma extremidade livre que invade e se une à extremidade livre da outra molécula de DNA. A invasão e a união de fita ocorrem em ambas as moléculas de DNA, criando dois DNA heteroduplex, cada um com uma fita original mais uma fita nova de outra molécula de DNA. O ponto no qual as fitas de nucleotídios passam de uma molécula de DNA para a outra é a junção de Holliday. A junção se desloca ao longo das moléculas em um processo chamado migração de ramificações. A troca de fitas de nucleotídios e a migração de ramificações produzem uma estrutura chamada intermediário de Holliday, que pode ser rompida em uma das duas formas. A ruptura no plano horizontal, seguida pela reunião das fitas, produz recombinantes não *crossing over*, no qual os genes de cada extremidade das moléculas são idênticos aos que havia originalmente (gene *A* com gene *B* e gene *a* com gene *b*). A ruptura no plano vertical, seguida pela reunião das fitas, produz recombinantes *crossing over*, no qual os genes da extremidade das moléculas são diferentes aos que havia originalmente (gene *A* com gene *b* e gene *a* com gene *B*).

Outra via para recombinação é iniciada com rupturas de fita dupla em uma das duas moléculas de DNA alinhadas (**Figura 12.21**). Nesse modelo, a remoção de alguns nucleotídios nas extremidades das fitas rompidas, seguida por invasão, deslocamento e replicação de fita, produz duas moléculas de DNA heteroduplex unidas por duas junções de Holliday. As moléculas interconectadas produzem o molde de ruptura de fita dupla que pode ser separado por ruptura e reunião das fitas de nucleotídios do mesmo modo que o intermediário de Holliday é separado no modelo de ruptura de fita dupla. A produção de moléculas *crossing over* ou não *crossing over* depende se a ruptura é no plano vertical ou no plano horizontal.

As evidências do modelo de ruptura de fita dupla provêm originalmente de resultados de cruzamentos genéticos na levedura que não poderiam ser explicados pelo modelo de Holliday. Observações subsequentes mostraram que as rupturas de fita dupla surgem na prófase I das leveduras, quando ocorre o *crossing over*, e que as cepas mutantes incapazes de formar rupturas de fita dupla não exibem recombinação meiótica. Embora exista evidência considerável que apoie o modelo de ruptura de fita dupla na levedura, a extensão que se aplique a outros organismos ainda é desconhecida.

Conceitos

A recombinação homóloga requer a formação de DNA heteroduplex, com uma fita de nucleotídios a partir de cada um dos dois cromossomos homólogos. No modelo de Holliday, a recombinação homóloga é obtida por meio de uma ruptura de fita única no DNA, deslocamento de fita e migração de ramificações. No modelo de ruptura de fita dupla, a recombinação é obtida por meio de rupturas de fita dupla, deslocamento de fita e migração de ramificações.

Checagem dos conceitos 11

Por que a recombinação é importante?

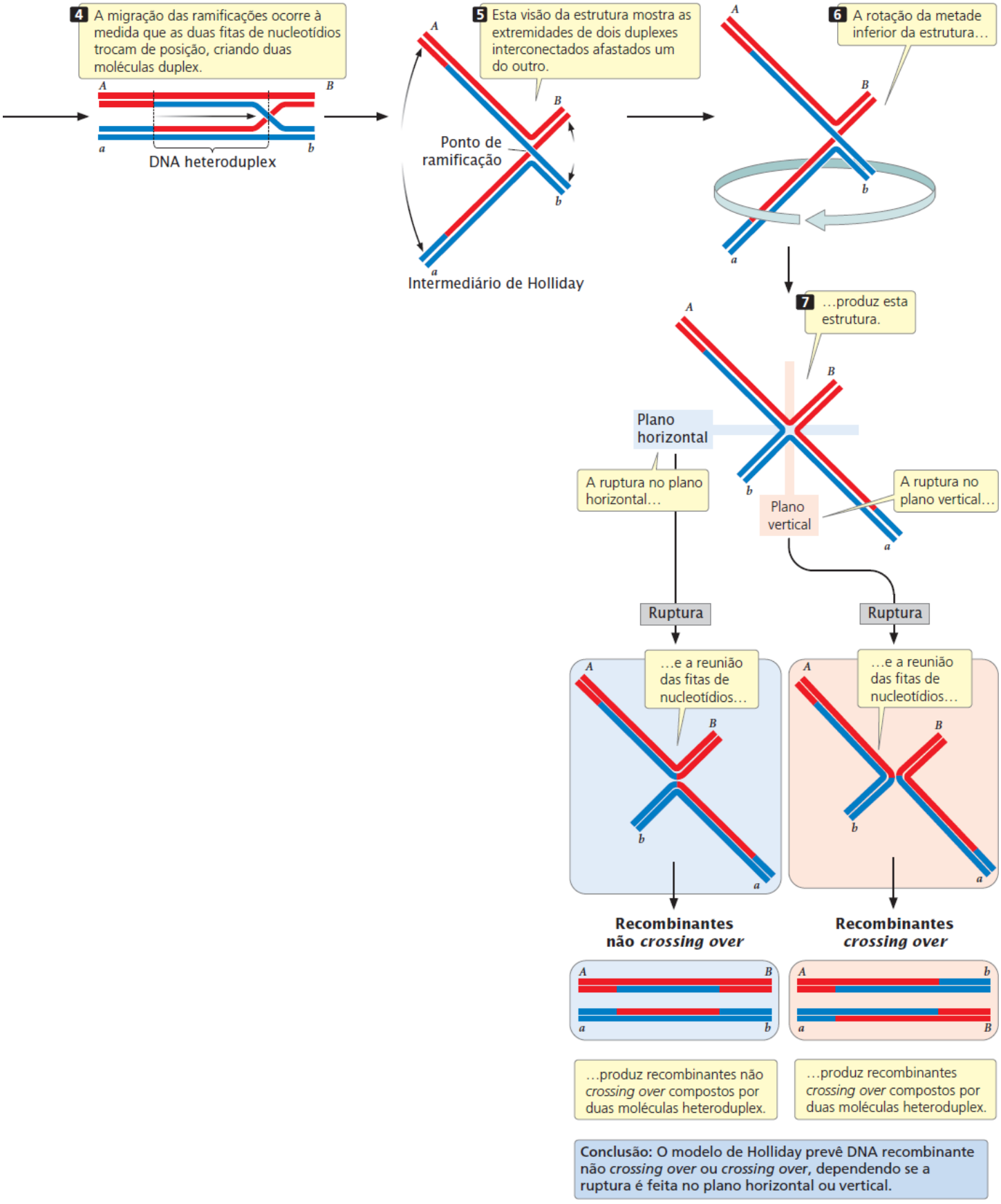


Figura 12.20 O modelo de Holliday da recombinação homóloga. Neste modelo, a recombinação ocorre por meio de uma ruptura de fita única em cada DNA duplex, deslocamento de fita, migração de ramo e resolução de uma junção de Holliday única.

Enzimas necessárias à recombinação

A recombinação entre moléculas de DNA requer o desenrolamento das hélices de DNA, a clivagem das fitas de nucleotídios, invasão de fita e migração de ramificações, seguida por clivagem adicional da fita e união para remover as junções de Holliday. Boa parte do que sabemos sobre esses processos surgiu a partir de estudos da troca de genes na *E. coli*. Embora as bactérias não sofram meiose, elas têm um tipo de reprodução sexuada (conjugação), na qual uma bactéria doa seu cromossomo para outra

(discutido com mais detalhes no *Capítulo 9*). Após a conjugação, a bactéria receptora tem dois cromossomos, que podem sofrer recombinação homóloga. Os geneticistas isolaram cepas mutantes de *E. coli* que não apresentam recombinação; o estudo dessas cepas levou à identificação dos genes e das proteínas que participam da recombinação bacteriana, revelando diversas vias pelas quais ela pode ser feita.

Três genes com papéis críticos na recombinação na *E. coli* são *recB*, *recC* e *recD*, responsáveis pela codificação de três polipeptídios que, juntos, formam a proteína RecBCD. Essa proteína desenrola o DNA de fita dupla e é capaz de romper fitas de nucleotídios. O gene *recA* codifica a proteína RecA; essa proteína possibilita que uma fita única invada uma hélice de DNA e o subsequente deslocamento de uma das fitas originais. Nos eucariotos, a formação e a migração de ramificações das estruturas de Holliday são facilitadas pela enzima Rad51.

Na *E. coli*, os genes *ruvA* e *ruvB* codificam proteínas que catalisam a migração da ramificação, e o gene *ruvC* produz uma proteína, chamada resolvase, que rompe as estruturas de Holliday. A ruptura e a resolução das estruturas de Holliday nos eucariotos são feitas por uma enzima análoga chamada GEN1. As proteínas ligadoras de fita única, DNA ligase, DNA polimerases e DNA girase também participam em vários tipos de recombinação, além de suas funções na replicação do DNA.

Conceitos

Várias proteínas têm papéis na recombinação, incluindo RecA, RecBCD, RuvA, RuvB, resolvase, proteínas ligadoras de DNA de fita única, ligase, DNA polimerases e girase.

Checagem dos conceitos 12

Qual é a função da resolvase na recombinação?

a. Desenrola o DNA de fita dupla.
b. Possibilita que uma fita de DNA única invada uma hélice de DNA.
c. Desloca uma das fitas de DNA originais durante a migração de ramificações.
d. Rompe a estrutura de Holliday.

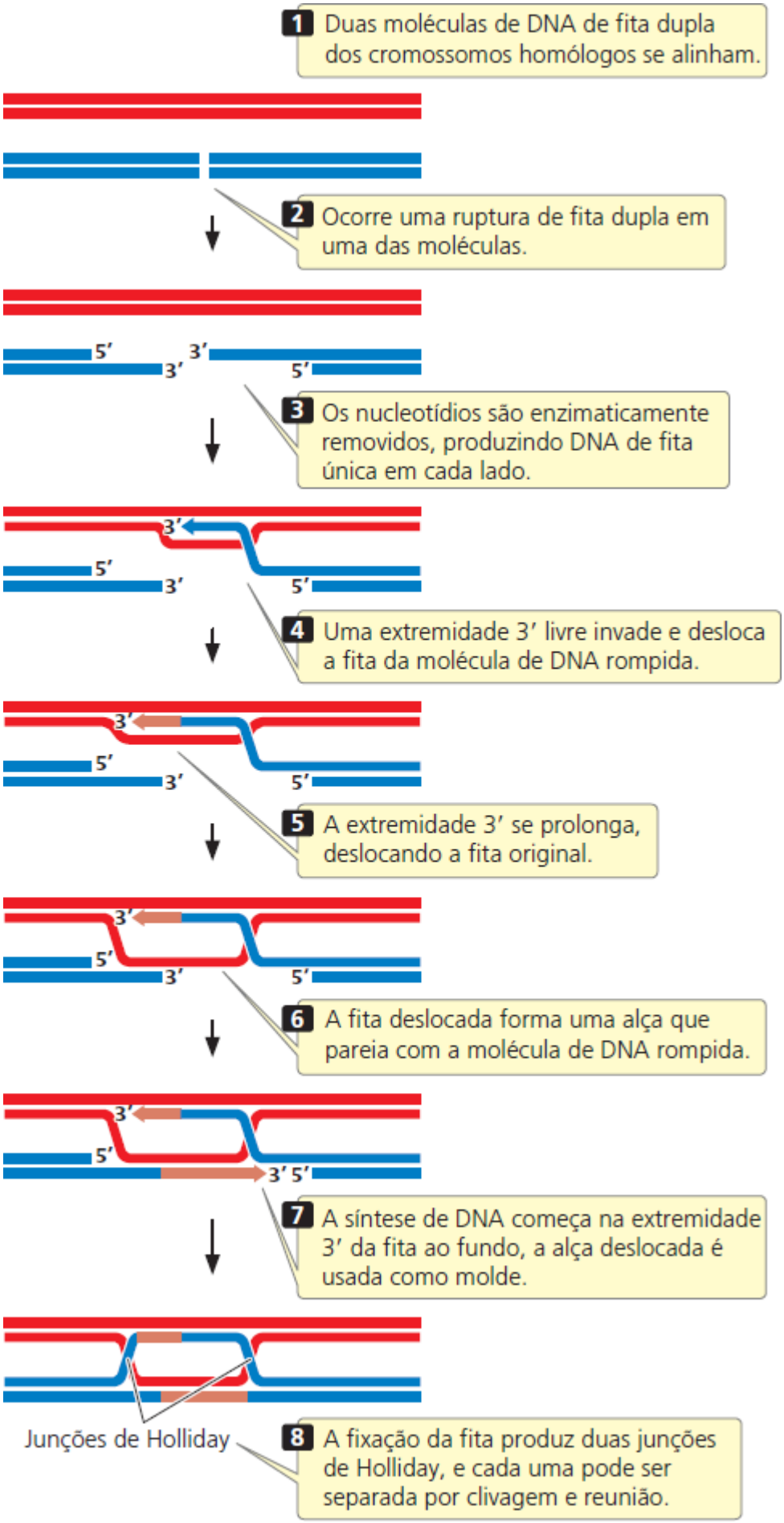


Figura 12.21 O modelo de ruptura de fita dupla de recombinação. Neste modelo, a recombinação ocorre por meio de ruptura de fita dupla em um DNA duplex, deslocamento de fita, síntese de DNA e resolução das duas junções de Holliday.

Conversão gênica

Como observamos, a recombinação homóloga é o mecanismo que produz o *crossing over*. Ela também é responsável por um fenômeno relacionado conhecido com **conversão gênica**, um processo de troca genética não recíproca que pode produzir razões anormais de gametas após a meiose. Por exemplo, um organismo individual com genótipo *Aa* deve produzir 1/2 gametas *A* e 1/2 gametas *a*. Algumas vezes, entretanto, a meiose em um indivíduo *Aa* produz 3/4 *A* e 1/4 *a* ou 1/4 *A* e 3/4 *a*. A conversão de genes surge a partir da formação heteroduplex que ocorre na recombinação. Durante a formação heteroduplex, uma molécula de DNA de fita única de um cromossomo pareia com uma molécula de DNA de fita única de outro cromossomo. Se as duas fitas em um heteroduplex surgem a partir de cromossomos com alelos diferentes, ocorre um pareamento errado das bases no DNA heteroduplex (**Figura 12.22**). Tais pareamentos errados são reparados pela célula. Os mecanismos de reparo frequentemente retiram os nucleotídios em uma das fitas e os substituem com DNA novo ao usar a fita complementar como molde. Uma cópia de um alelo pode ser convertida em outro alelo, levando a um evento de conversão gênica (ver Figura 12.22), dependendo de qual fita serve como molde.

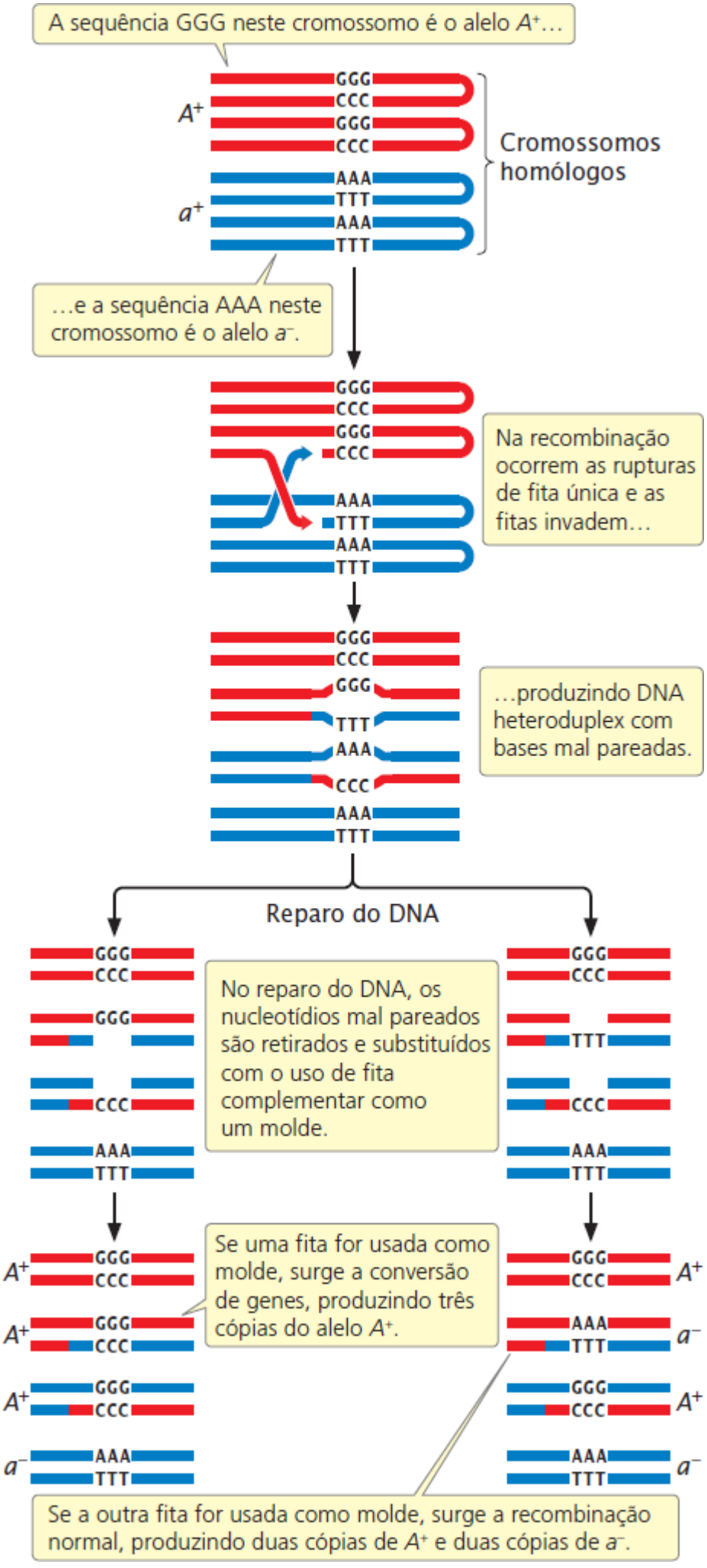

A sequência GGG neste cromossomo é o alelo A+...
A+
GGG
CCC
GGG
CCC
Cromossomos homólogos
a+
AAA
TTT
AAA
TTT
...e a sequência AAA neste cromossomo é o alelo a–.
Na recombinação ocorrem as rupturas de fita única e as fitas invadem...
...produzindo DNA heteroduplex com bases mal pareadas.
Reparo do DNA
No reparo do DNA, os nucleotídios mal pareados são retirados e substituídos com o uso de fita complementar como um molde.
Se uma fita for usada como molde, surge a conversão de genes, produzindo três cópias do alelo A+.
Se a outra fita for usada como molde, surge a recombinação normal, produzindo duas cópias de A+ e duas cópias de a–.

Figura 12.22 A conversão de genes ocorre por meio do reparo de bases mal pareadas no DNA heteroduplex.

Resumo dos conceitos

- A replicação é semiconservativa. Duas fitas de nucleotídios de DNA se separam e cada uma serve como molde para a síntese de uma nova fita
- Na replicação teta do DNA, as duas fitas de nucleotídios de uma molécula de DNA circular se desenrolam, criando uma bolha de replicação; dentro de cada bolha de replicação o DNA é normalmente sintetizado em ambas as fitas e em ambas as forquilhas de replicação, produzindo duas moléculas de DNA circular
- A replicação em círculo rolante é iniciada por uma ruptura em uma fita do DNA circular, que produz um grupo 3′-OH ao qual novos nucleotídios são adicionados enquanto a extremidade 5′ da fita rompida é deslocada do círculo
- O DNA eucariótico linear tem muitas origens de replicação. O desenrolamento e a replicação ocorrem em ambos os moldes nas extremidades da bolha de replicação até os replicons adjacentes se encontrarem, gerando duas moléculas de DNA lineares
- Toda a síntese de DNA é feita no sentido 5′ → 3′. Como as duas fitas de nucleotídios do DNA são antiparalelas, a replicação ocorre de maneira contínua em uma fita (a fita líder) e descontínua na outra (a fita tardia)
- A replicação começa quando uma proteína de iniciação se liga à origem da replicação e desenrola um trecho curto de DNA ao qual a DNA helicase se fixa. A DNA helicase desenrola o DNA na forquilha de replicação, as proteínas ligadoras de DNA de fita única se ligam a fitas únicas de nucleotídios para evitar estruturas secundárias e a DNA girase (uma topoisomerase) remove a torção à frente da forquilha de replicação que é gerada pelo desenrolamento
- Durante a replicação, a primase sintetiza *primers* curtos de nucleotídios de RNA, fornecendo um grupo 3′-OH ao qual a DNA polimerase pode adicionar nucleotídios de DNA
- A DNA polimerase adiciona novos nucleotídios à extremidade 3′ de uma fita de nucleotídios crescente. As bactérias têm duas DNA polimerases com papéis principais na replicação: a DNA polimerase III, que sintetiza DNA novo na fita líder e na fita tardia, e a DNA polimerase I, que remove e substitui os *primers*
- A DNA ligase fecha as rupturas que permanecem nos esqueletos de açúcar-fosfato, quando os *primers* de RNA são substituídos por nucleotídios de DNA
- Vários mecanismos garantem a alta taxa de exatidão na replicação, incluindo a seleção precisa de nucleotídios, a revisão e o reparo de pareamento errado
- A replicação precisa em múltiplas origens nos eucariotos é garantida por um fator de licenciamento que precisa se fixar a uma origem antes que a replicação possa começar
- Os nucleossomos eucarióticos são rapidamente montados em novas moléculas de DNA; os nucleossomos recém-montados são uma mistura aleatória de histonas novas e antigas
- As extremidades das moléculas de DNA eucarióticas lineares são replicadas pela enzima telomerase
- A recombinação homóloga ocorre por meio de rupturas nas fitas de nucleotídios, alinhamento de segmentos homólogos de DNA e reunião das fitas. Várias enzimas e proteínas são necessárias para a recombinação homóloga
- A conversão gênica é a troca genética não recíproca e produz razões anormais de gametas.

Termos importantes

Bolha de replicação
Centrifugação por gradiente de densidade de equilíbrio
Conversão gênica
DNA girase
DNA helicase
DNA heteroduplex
DNA ligase
DNA polimerase
DNA polimerase I
DNA polimerase III
DNA polimerase translesão
DNA polimerase α
DNA polimerase δ
DNA polimerase ε
Fator de licenciamento de replicação

Fita líder
Fita tardia
Forquilha de replicação
Fragmento de Okazaki
Junção de Holliday
Origem de replicação
Overhang G
Primase
Primer
Proteína de iniciação
Proteína ligadora de DNA de fita única (SSB)
Recombinação homóloga
Reparo de pareamento errado
Replicação bidirecional
Replicação contínua
Replicação descontínua
Replicação por círculo rolante
Replicação semiconservativa
Replicação teta
Replicon
Revisão
Sequência de replicação autônoma (ARS)
Telomerase

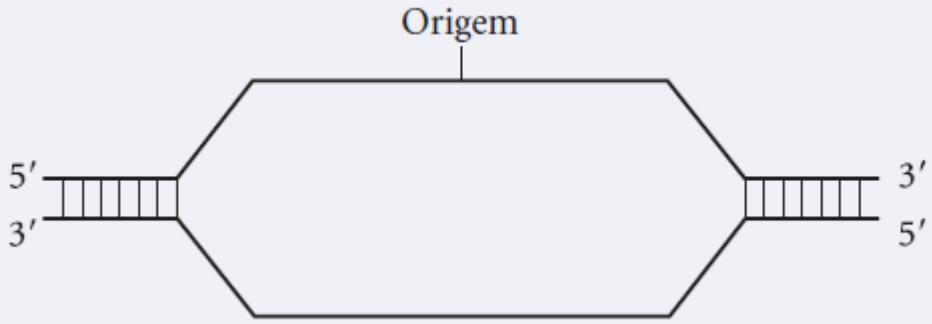
Respostas da Checagem dos conceitos

1. Duas bandas
2. b.
3. c.
4. Proteína de iniciação, helicase, proteína ligadora de DNA de fita única, DNA girase.
5. c.
6. b.
7. c.
8. O tamanho dos genomas eucarióticos, a estrutura linear dos cromossomos eucarióticos e a associação do DNA com as proteínas histonas.
9. Porque as DNA polimerases suscetíveis a erros conseguem contornar as lesões na hélice do DNA que paralisam as DNA polimerases exatas, de alta velocidade.
10. c.
11. A recombinação é importante para a variação genética e para alguns tipos de reparo do DNA.
12. d.

Problemas desenvolvidos

Problema 1

O diagrama a seguir representa as fitas molde de uma bolha de replicação em uma molécula de DNA. Desenhe as fitas recém-sintetizadas e identifique a fita líder e a fitatardia.



Estratégias para a solução

De quais informações você precisa para solucionar o problema?
O diagrama acima com as fitas recém-sintetizadas desenhadas e a fita líder e a fita tardia marcadas.

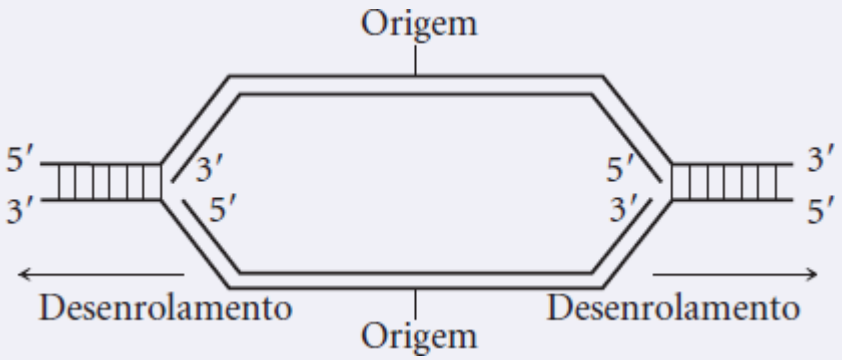
Quais informações são fornecidas para solucionar o problema?
Um diagrama do DNA molde com as extremidades 5′ e 3′ marcadas.

Para a solução deste problema, revise:
Sentido de replicação, na Seção 12.2, e Figura 12.10 C .

Etapas para a solução

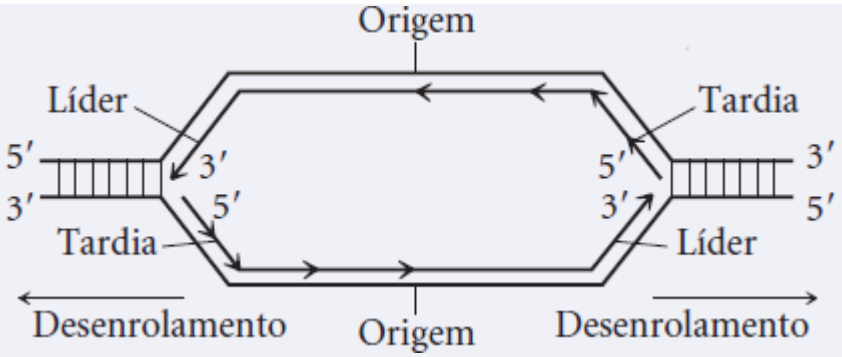
Para determinar a fita líder e a fita tardia, primeiro observe qual extremidade de cada fita molde é 5′ e qual é 3′. Com um lápis, desenhe nas fitas que estão sendo sintetizadas nesses moldes e identifique suas extremidades 5′ e 3′.

Lembrete: As duas fitas de DNA são antiparalelas, então a fita recém-sintetizada deve ter a polaridade oposta (direção) em relação à fita molde.



A seguir, determine a direção da replicação para cada nova fita, que deve ser 5′ → 3′. Você pode desenhar setas nas fitas novas para indicar a direção da replicação. Após ter estabelecido a direção da replicação para cada fita, observe cada forquilha e determine se a direção da replicação para uma fita é a mesma que a direção do desenrolamento. A fita na qual a replicação está na mesma direção do desenrolamento é a fita líder. A fita na qual a replicação está na direção oposta do desenrolamento é a fita tardia.

Lembrete: A síntese de DNA é sempre de 5′ para 3′.



Dica: Cada forquilha de replicação tem uma fita líder e uma fita tardia.

Problema 2

Considere o experimento feito por Meselson e Stahl no qual eles usaram ^{14}N e ^{15}N nas culturas de *E. coli* e a centrifugação por gradiente de densidade de equilíbrio. Desenhe imagens para representar as bandas produzidas pelo DNA bacteriano no tubo de gradiente de densidade antes de trocar para o meio contendo ^{14}N e após um, dois ou três ciclos de replicação depois da troca para o meio contendo ^{14}N. Use um conjunto separado de desenhos para mostrar as bandas que apareceriam se a replicação fosse (a) semiconservativa; (b) conservativa; (c) dispersiva.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

Desenhos que representem as bandas produzidas pelo DNA bacteriano em tubos de gradiente de densidade antes da troca para meio contendo ^{14}N e após um, dois e três ciclos de replicação seguidos da troca para o meio contendo ^{14}N; assim, você deve ter os desenhos dos quatro tubos para cada modelo de replicação. Você precisa de um conjunto separado de desenhos para a replicação semiconservativa, conservativa e dispersiva.

Quais informações são fornecidas para solucionar o problema?

- O DNA bacteriano foi originalmente marcado com ^{15}N e, então, as bactérias foram transferidas para um meio com ^{14}N (ver discussão do experimento nas pp. 294-295)
- O DNA original terá ^{15}N; o DNA recém-sintetizado terá ^{14}N
- A centrifugação por gradiente de densidade foi realizada antes da troca para ^{14}N e após um, dois e três ciclos de replicação depois da troca.

Para a solução deste problema, revise:

Experimento de Meselson e Stahl, na Seção 12.2.

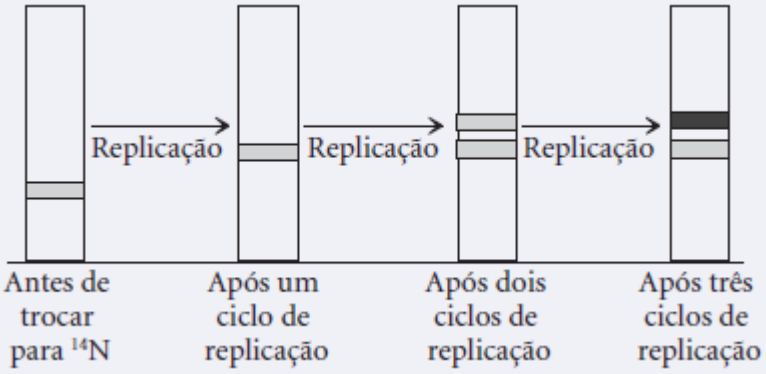
Etapas para a solução

O DNA marcado com ^{15}N será mais denso do que o marcado com ^{14}N; portanto, o DNA marcado com ^{15}N ficará no fundo do tubo de gradiente de densidade. Antes da troca para o meio com ^{14}N, todo o DNA da bactéria terá ^{15}N e produzirá uma única banda na parte inferior do tubo.

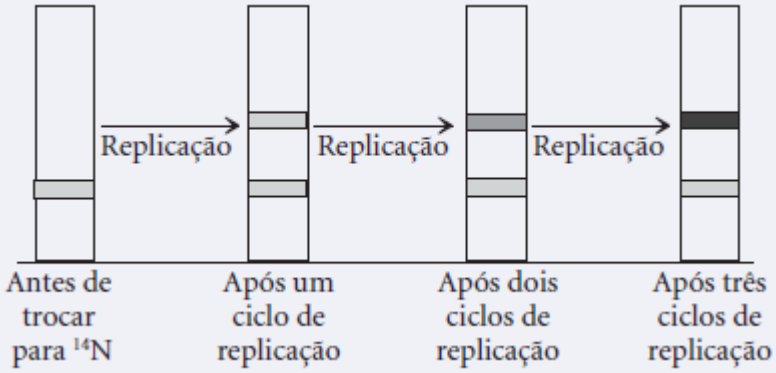
a. Na replicação semiconservativa, as duas fitas se separam e cada uma serve como molde no qual uma nova fita é sintetizada. Após um ciclo de replicação, a fita molde original de cada molécula terá ^{15}N, e a nova fita de cada molécula terá ^{14}N. Então, uma banda única aparecerá no meio do gradiente de densidade entre as posições esperadas do DNA com apenas ^{15}N e do DNA com apenas ^{14}N. No ciclo seguinte de replicação, as duas fitas se separam novamente e servem como moldes para novas fitas. Cada uma das novas fitas tem apenas ^{14}N, então algumas moléculas de DNA terão uma fita com o ^{15}N original e uma fita com o ^{14}N novo, enquanto outras moléculas terão duas fitas com ^{14}N. Essa marcação produzirá duas bandas, uma na posição intermediária e outra na porção superior do tubo. Ciclos adicionais de replicação produzirão quantidades crescentes de DNA com apenas ^{14}N; então, a banda superior ficará mais escura.

Dica: Revise a distribuição de DNA novo e antigo na replicação semiconservativa, conservativa e dispersiva na

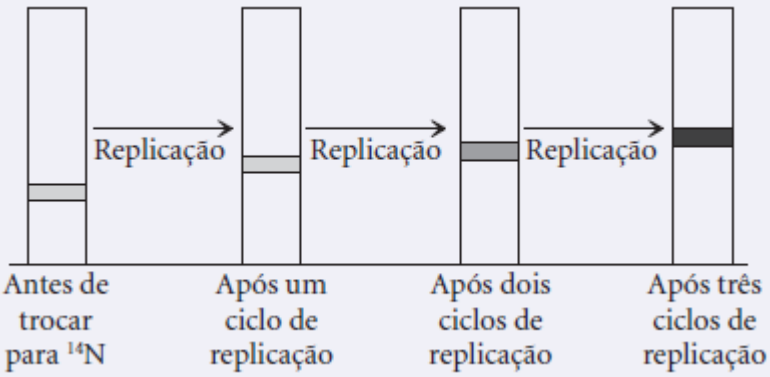
Figura 12.1 .



b. Na replicação conservativa, a molécula inteira serve como molde. Após um ciclo de replicação, algumas moléculas terão apenas ^{15}N, e outras terão apenas ^{14}N; então, deve haver duas bandas. Ciclos subsequentes de replicação aumentarão a fração de DNA consistindo em apenas ^{14}N novo; assim, a banda superior ficará mais escura. Entretanto, o DNA original com ^{15}N permanecerá, e as duas bandas estarão apresentadas.



c. Na replicação dispersiva, ambas as fitas de nucleotídios são rompidas em fragmentos que servem como moldes para a síntese de novo DNA. Os fragmentos são remontados em moléculas de DNA. Após um ciclo de replicação, todo o DNA deve ter aproximadamente metade de ^{15}N e metade de ^{14}N, produzindo uma banda única que está no meio entre as posições esperadas para o DNA marcado com ^{15}N e o DNA marcado com ^{14}N. Após outros ciclos de replicação, a proporção de ^{14}N em cada molécula aumenta; então, uma banda híbrida única permanece, mas sua posição no gradiente de densidade se desloca para cima. Também é esperado que a banda fique mais escura à medida que a quantidade total de DNA aumente.



Questões de compreensão

Seção 12.2

1. O que é replicação semiconservativa?
2. Como Meselson e Stahl demonstraram que a replicação na *E. coli* ocorre de modo semiconservativo?
3. Desenhe uma molécula de DNA submetida à replicação teta. Em seu desenho, identifique (a) origem, (b) polaridade (extremidades 5′ e 3′) de todas as fitas molde e fitas recém-sintetizadas, (c) fita líder e fita tardia, (d) fragmentos de Okazaki e (e) localização dos *primers*.
4. Desenhe uma molécula de DNA submetida à replicação em círculo rolante. Em seu desenho, identifique (a) origem, (b) polaridade (extremidades 5′ e 3′) de todas as fitas molde e fitas recém-sintetizadas, (c) fita líder e fita tardia, (d) fragmentos de Okazaki e (e) localização dos *primers*.
5. Desenhe uma molécula de DNA submetida à replicação eucariótica linear. Em seu desenho, identifique (a) origem, (b) polaridade (extremidades 5′ e 3′) de todas as fitas molde e fitas recém-sintetizadas, (c) fita líder e fita tardia, (d) fragmentos de Okazaki e (e) localização dos *primers*.
6. Quais são as três principais exigências da replicação?
7. Que substratos são usados na reação de síntese de DNA?

Seção 12.3

8. Liste as diferentes proteínas e enzimas envolvidas na replicação bacteriana. Indique a função de cada uma nesse processo.
9. Por que a DNA girase é necessária para a replicação?
10. Quais são as semelhanças e diferenças existentes nas atividades enzimáticas das DNA polimerases I e III? Qual é a função de cada DNA polimerase nas células bacterianas?
11. Por que a primase é necessária para a replicação?
12. Quais são os três mecanismos que garantem a precisão da replicação nas bactérias?

Seção 12.4

13. Como o licenciamento da replicação garante que o DNA seja replicado apenas uma vez em cada origem por ciclo de célula eucariótica?
14. Quais são as semelhanças e diferenças na replicação eucariótica e na replicação bacteriana?
15. O que é o problema da extremidade do cromossomo para a replicação linear? Por que, na ausência da telomerase, as extremidades dos cromossomos se tornam progressivamente mais curtas cada vez que o DNA é replicado?
16. Descreva em palavras e desenhos como os telômeros nas extremidades dos cromossomos eucarióticos são replicados.

Seção 12.5

17. Explique como o tipo de ruptura do intermediário de Holliday produz recombinantes não *crossing over* e recombinantes.
18. Quais são algumas das enzimas envolvidas na recombinação na *E. coli* e quais são seus papéis?
19. O que é conversão gênica? Como ela surge?

Questões e problemas aplicados

Seção 12.2

20. Imagine que um cientista do futuro explore um planeta distante e descubra uma nova forma de ácido nucleico de fita dupla. Quando esse ácido nucleico é exposto à DNA polimerase da *E. coli*, a replicação ocorre de maneira contínua em ambas as fitas. Que conclusão você pode tirar sobre a estrutura desse novo ácido nucleico?
21. O fósforo é necessário para sintetizar os trifosfatos de desoxirribonucleosídio usados na replicação do DNA. Um geneticista cultiva *E. coli* em um meio com fósforo não radioativo por muitas gerações. Uma amostra das bactérias é então transferida para um meio com o isótopo radioativo do fósforo (^{32}P). Amostras das bactérias são removidas imediatamente após a transferência e após um e dois ciclos de replicação. Assuma que o DNA recém-sintetizado tem ^{32}P e o DNA original tem fósforo não radioativo. Qual será a distribuição de radioatividade no DNA da bactéria em cada amostra? A radioatividade será detectada em uma, ambas ou nenhuma das fitas de DNA?

*22. Uma linhagem de células de camundongos é cultivada por muitas gerações em um meio com ^{15}N. As células na fase G_1 são trocadas para um novo meio com ^{14}N. Desenhe um par de cromossomos homólogos a partir dessas células nos estágios a seguir, mostrando as duas fitas de moléculas de DNA encontradas nos cromossomos. Use diferentes cores para representar as fitas com ^{14}N e ^{15}N.
    a. Células em G_1 antes da troca para meio com ^{14}N.
    b. Células em G_2 após a troca para meio com ^{14}N.
    c. Células na anáfase da mitose, após a troca para meio com ^{14}N.

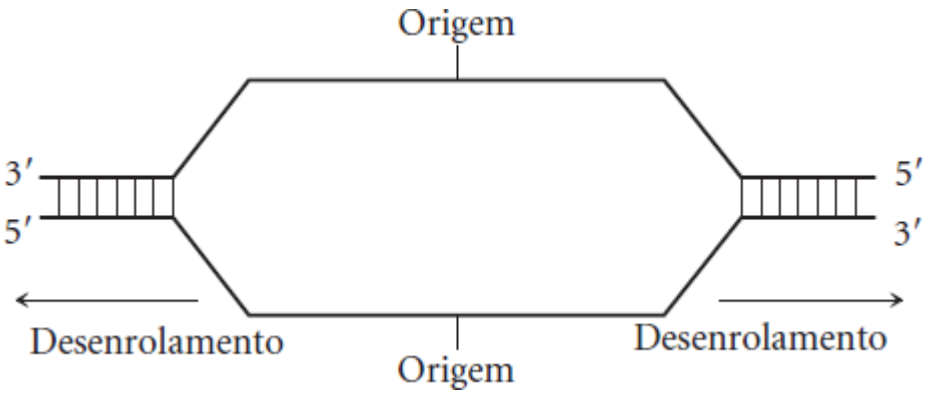
d. Células na metáfase I da meiose, após a troca para meio com ^{14}N.
e. Células na metáfase anáfase II da meiose, após a troca para meio com ^{14}N.

*23. Uma molécula circular de DNA tem 1 milhão de pares de base. Se a taxa de síntese de DNA em uma forquilha de replicação for 100.000 nucleotídios por minuto, quanto tempo a replicação teta precisa para replicar completamente a molécula, considerando que a replicação teta é bidirecional? Quanto tempo vai durar a replicação por ciclo rolante para esse cromossomo circular? Ignore a replicação da fita deslocada na replicação por círculo rolante.

24. Uma bactéria sintetiza DNA em cada forquilha de replicação em uma taxa de 1.000 nucleotídios por segundo. Se essa bactéria replicar completamente seu cromossomo circular pela replicação teta em 30 minutos, quantos pares de base de DNA terá seu cromossomo?

Seção 12.3

*25. O diagrama a seguir representa uma molécula de DNA durante sua replicação. Desenhe as fitas do DNA recém-sintetizado e identifique (a) a polaridade das fitas recém-sintetizadas, (b) a fita líder e a fita tardia, (c) os fragmentos de Okazaki e (d) os *primers* de RNA.



26. Na Figura 12.8, identifique qual fita é a líder e qual é a tardia.

*27. Qual seria o efeito na replicação do DNA de mutações que destroem cada uma das atividades da DNA polimerase I?
a. Atividade exonuclease 3′ → 5′.
b. Atividade exonuclease 5′ → 3′.
c. Atividade polimerase 5′ → 3′.

28. Qual das DNA polimerases apresentadas no Quadro 12.3 é responsável pela revisão?

29. Como a replicação do DNA seria afetada em uma célula bacteriana que não tem DNA girase?

*30. Qual seria o efeito na replicação teta se o gene para primase fosse um mutante de modo que nenhuma primase fosse produzida? Na replicação por círculo rolante?

31. As DNA polimerases não são capazes de iniciar a replicação, embora a primase e a RNA polimerases possam. Alguns geneticistas especularam que a incapacidade da DNA polimerase em iniciar a replicação esteja relacionada com sua função de revisão. Essa hipótese defende que a revisão é essencial para a transmissão fidedigna da informação genética e que, como as DNA polimerases evoluíram a capacidade de revisar, elas não podem iniciar a síntese do DNA. Explique por que as funções de revisão e iniciação na mesma enzima seriam incompatíveis.

Seção 12.4

ANÁLISE DE DADOS

32. Marina Melixetian e seus colegas suprimiram a expressão da proteína Geminina nas células humanas ao tratar as células com pequenos RNA interferentes (siRNAs) complementares ao RNA mensageiro da Geminina (M. Melixetian *et al*. 2004. *Journal of Cell Biology* 165:473-482). (Pequenos RNAs interferentes formam um complexo com proteínas e pareiam com as sequências complementares nos mRNAs; o complexo rompe o mRNA e não ocorre tradução desse mRNA; ver *Capítulo 14*.) Quarenta e oito horas após o tratamento com siRNA, as células sem Geminina estavam maiores e apresentavam um único núcleo gigante. A análise do conteúdo de DNA mostrou que muitas dessas células sem Geminina eram 4*n* ou mais. Explique esses resultados.

*33. Quais resultados seriam esperados no experimento destacado na Figura 12.17 se, durante a replicação, todas as histonas originais permanecessem em uma fita de DNA e as histonas novas ficassem ligadas na outra fita?

34. Vários cientistas que estudam tratamento do câncer têm interesse na telomerase. Por quê? Como as terapias

medicamentosas para câncer direcionadas para telomerase poderiam atuar?

*35. A enzima telomerase é parte proteína e parte RNA. Qual seria o efeito mais provável de uma grande deleção no gene que codifica a parte RNA da telomerase? Como a função da telomerase seria afetada?

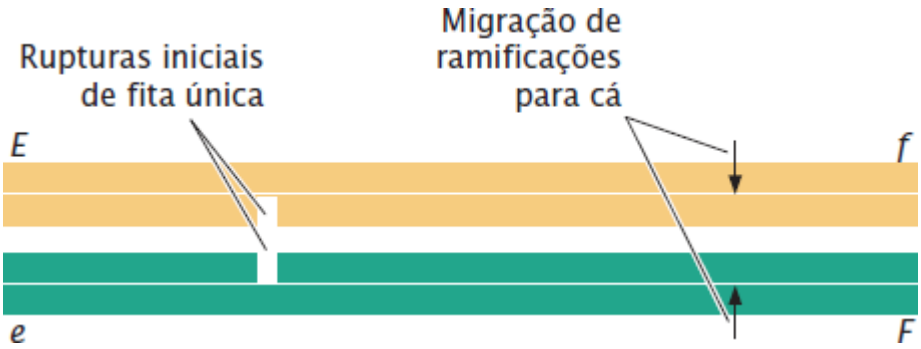
ANÁLISE DE DADOS

36. A disqueratose congênita (DKC) é um raro distúrbio genético caracterizado por unhas dos dedos das mãos e pigmentação da pele anormais, formação de placas brancas na língua e bochecha e insuficiência progressiva da medula óssea. Uma forma autossômica dominante de DKC surge a partir de mutações no gene que codifica o componente RNA da telomerase. Tom Vulliamy e seus colegas examinaram 15 famílias com DKC dominante autossômica (T. Vulliamy *et al*. 2004. *Nature Genetics* 36:447-449). Eles observaram que a idade mediana de início da DKC nos genitores era 37 anos, enquanto nos filhos dessas pessoas era 14,5 anos. Assim, a DKC nessas famílias surge em idades progressivamente precoces nas gerações sucessivas, um fenômeno conhecido como antecipação (ver *Capítulo 5*). Os pesquisadores mediram o comprimento do telômero dos membros dessas famílias; as medidas são fornecidas na tabela a seguir. O comprimento do telômero reduz com a idade; então, esse comprimento foi ajustado pela idade. Observe que o comprimento ajustado pela idade de todos os membros dessas famílias é negativo, indicando que seus telômeros são menores que o normal. Para o comprimento do telômero ajustado pela idade, quanto mais negativo o número, mais curto o telômero.

Comprimento do telômero do genitor	Comprimento do telômero do filho
−4,7	−6,1
	−6,6
	−6,0
−3,9	−0,6
−1,4	−2,2
−5,2	−5,4
−2,2	−3,6
−4,4	−2,0
−4,3	−6,8
−5,0	−3,8
−5,3	−6,4
−0,6	−2,5
−1,3	−5,1
	−3,9
−4,2	−5,9

a. Como o comprimento do telômero dos genitores se compara com o dos filhos? (Dica: Calcule o comprimento médio do telômero de todos os pais e o comprimento médio do telômero de todas as crianças.)
b. Explique por que os telômeros de pessoas com DKC são mais curtos que o normal.
c. Explique por que a DKC surge em idade mais precoce nas gerações seguintes.

37. Um indivíduo é heterozigoto em dois *loci* (*Ee Ff*) e os genes estão em repulsão (ver *Capítulo 7*). Considere que as rupturas de fita única e a migração de ramificações ocorrem nas posições apresentadas a seguir. Usando diferentes cores para representar os dois cromossomos homólogos, desenhe as moléculas de DNA recombinantes não *crossing over* e recombinantes *crossing over* que surgem a partir da recombinação homóloga (dica: ver Figura 12.20).



Seção 12.3

38. Uma mutação condicional expressa seu fenótipo mutante apenas em algumas condições (as condições restritivas) e seu fenótipo normal em outras condições (as condições permissivas). Um tipo de mutação condicional é uma mutação sensível à temperatura, que expressa o fenótipo mutante apenas em algumas temperaturas.

    Foram isoladas cepas de *E. coli* com mutações termossensíveis nos genes que codificam diferentes componentes do maquinário de replicação. Em cada uma dessas cepas, a proteína produzida pelo gene mutante não é funcional nas condições restritivas. Essas cepas crescem nas condições permissivas e então, abruptamente, mudam para a condição restritiva. Após um ciclo de replicação em condição restritiva, o DNA de cada cepa é isolado e analisado. Quais características você esperaria encontrar no DNA isolado de cada cepa com uma mutação sensível à temperatura no seu gene que codifica as seguintes proteínas?

    a. DNA ligase.
    b. DNA polimerase I.
    c. DNA polimerase III.
    d. Primase.
    e. Proteína de iniciação.

Seção 12.4

39. As DNA topoisomerases têm papéis importantes na replicação do DNA e no superenrolamento (ver *Capítulo 11*). Essas enzimas também são alvos de alguns fármacos anticâncer. Eric Nelson e seus colegas estudaram m-AMSA, um dos compostos anticâncer que atua nas enzimas topoisomerase (E. M. Nelson, K. M. Tewey, L. F. Liu. 1984. *Proceedings of the National Academy of Sciences* 81:1361-1365). Eles descobriram que m-AMSA estabiliza um intermediário produzido no curso de ação da topoisomerase. O intermediário é a topoisomerase ligada às extremidades rompidas do DNA. As rupturas no DNA que são produzidas por compostos anticâncer como m-ASMA inibem a replicação do DNA celular e impedem que as células do câncer proliferem. Proponha um mecanismo sobre como m-ASMA e outros agentes anticâncer direcionados para as enzimas topoisomerase envolvidas na replicação podem causar rupturas no DNA e rearranjos cromossômicos.

*40. A regulação da replicação é essencial para a estabilidade genômica e, normalmente, o DNA é replicado apenas uma vez em cada ciclo celular eucariótico (na fase S). As células normais produzem a proteína A, cuja concentração aumenta na fase S. Nas células que têm uma cópia mutante do gene para a proteína A, a proteína não é funcional e a replicação ocorre de modo contínuo durante todo o ciclo celular, resultando em células que podem ter 50 vezes a quantidade normal de DNA. A proteína B é encontrada normalmente em G_1, mas desaparece do núcleo celular na fase S. Nas células com uma cópia mutante do gene para proteína A, os níveis de proteína B não desaparecem na fase S e, pelo contrário, permanecem elevados durante todo o ciclo celular. Não ocorre replicação quando o gene para a proteína B é mutante.

    Proponha um mecanismo que explique como as proteínas A e B poderiam regular normalmente a replicação de modo que cada célula receba uma quantidade adequada de DNA. Explique como a mutação desses genes produz os efeitos descritos aqui.

13 Transcrição

Envenenamento por cicuta-verde

No dia 8 de novembro de 2009, Tomasa, de 31 anos, estava caminhando na trilha de Lodi Lake, a oeste de São Francisco, com o marido e o primo, quando encontrou cogumelos brancos e grandes que se pareciam muito com os cogumelos comestíveis que eles experimentaram no México. Eles pegaram os cogumelos e levaram para casa, cozinharam e comeram no jantar. Após algumas horas, Tomasa e sua família passaram mal e foram para o hospital. Eles foram transferidos para a unidade de cuidados críticos no California Pacific Medical Center em São Francisco, onde Tomasa morreu de insuficiência hepática após 3 semanas. O marido sobreviveu após um longo período de hospitalização, e o sobrinho precisou de um transplante de fígado.

Os cogumelos que eles consumiram eram *Amanita phalloides*, também conhecido como cicuta-verde. Uma única cicuta-verde tem toxina suficiente para matar um humano adulto. A taxa de morte entre os que consumiram esse cogumelo é de 22%; em crianças com menos de 10 anos, a taxa é maior que 50%. Aparentemente, esse tipo de cogumelo está se espalhando na Califórnia, levando a um recente aumento no número de envenenamentos.

Esse envenenamento é traiçoeiro. Os sintomas gastrintestinais – dor abdominal, cólicas, vômito, diarreia – começam entre 6 e 12 horas após a ingestão, mas melhoram após algumas horas, e o paciente apresenta melhora do quadro. Por causa dessa remissão inicial, o envenenamento não é levado a sério até que seja tarde demais para esvaziar o estômago e remover a toxina do corpo. Os sintomas graves começam após 1 ou 2 dias. As células hepáticas morrem, causando dano permanente ao fígado e a morte do indivíduo em alguns dias. Não existe tratamento efetivo, além do transplante de fígado, para substituir o órgão lesionado.

Como esse cogumelo mata? Sua toxina letal, contida nos corpos frutíferos que produzem os esporos reprodutivos, é a proteína α-amanitina, composta por um peptídio curto de oito aminoácidos que forma uma alça circular. A α-amanitina é um inibidor potente da RNA polimerase II, a enzima que transcreve os genes codificadores das proteínas nos eucariotos. A RNA polimerase II se liga aos genes e sintetiza moléculas de RNA complementares ao molde de DNA. No processo de transcrição, a RNA polimerase desloca o molde de DNA, adicionando um nucleotídio por vez à cadeia crescente de RNA. A α-amanitina se liga à RNA polimerase e pressiona as partes móveis da enzima, comprometendo sua capacidade de se mover ao longo do molde de DNA. Quando há α-amanitina, a síntese do RNA cai da sua faixa normal de centenas de milhares de nucleotídios por minuto para apenas alguns nucleotídios por minuto. Os resultados são catastróficos. Sem a transcrição, a síntese de proteínas – necessária ao funcionamento da célula– para, e então a célula morre. O fígado, onde a toxina se acumula, é lesionado de modo irreparável e para de funcionar. Nos casos graves, o paciente morre.

O envenenamento por cicuta-verde ilustra a extrema importância da transcrição e o papel crucial que a RNA polimerase desempenha nesse processo. Este capítulo discute o processo de transcrição – a primeira etapa no dogma central, a via de transferência de informação do DNA (genótipo) para a proteína (fenótipo). A transcrição é um processo complexo que requer precursores aos nucleotídios de RNA, um molde de DNA e vários componentes proteicos. À medida que examinamos os estágios da transcrição, tentamos manter todos os detalhes em perspectiva e nos concentrar em compreender como eles estão relacionados com o objetivo geral da transcrição – a síntese seletiva de uma molécula de RNA.

Este capítulo começa com uma rápida revisão da estrutura do RNA e uma discussão sobre as diferentes classes de RNA. Então analisamos os principais componentes necessários para a transcrição. Finalmente, exploramos o processo de transcrição. Em vários pontos do texto, pausaremos para analisar alguns princípios gerais.

13.1 O RNA, composto por uma fita única de ribonucleotídios, participa em várias funções celulares

Antes de começarmos nosso estudo sobre transcrição, analisaremos a importância passada e atual do RNA, revisaremos sua estrutura e examinaremos alguns dos tipos diferentes de moléculas de RNA.

Um mundo primitivo de RNA

A vida precisa de duas funções básicas. Primeiro, os organismos vivos têm de ser capazes de armazenar e transmitir de forma fidedigna a informação genética durante a reprodução. Segundo, precisam ter a capacidade de catalisar as transformações químicas que conduzem os processos da vida. Por muito tempo, acreditava-se que as funções do armazenamento de informações e transformação química fossem controladas por dois tipos de moléculas completamente diferentes: a informação genética seria armazenada nos ácidos nucleicos, enquanto as transformações químicas seriam catalisadas pelas enzimas. Essa dicotomia bioquímica criou um dilema. O que veio primeiro: a proteína ou o ácido nucleico? Se os ácidos nucleicos carreiam as instruções de codificação para as proteínas, de que maneira elas poderiam ser geradas sem eles? Como os ácidos nucleicos não são capazes de se autocopiar, como eles poderiam ser gerados sem as proteínas? Se o DNA e as proteínas precisam um do outro, como a vida teria começado?

Esse aparente paradoxo foi respondido em 1981 quando Thomas Cech e seus colegas descobriram que o RNA pode servir como catalisador biológico. Eles descobriram que o RNA do protozoário *Tetrahymena thermophila* consegue excisar 400 nucleotídios de seu RNA na ausência de qualquer proteína. Atualmente, foram descobertos outros exemplos de RNA catalíticos em diferentes tipos de células. Chamadas **ribozimas**, essas moléculas de RNA catalíticas conseguem cortar partes de sua própria sequência, unir algumas moléculas de RNA, replicar outras e até catalisar a formação de ligações de peptídios entre os aminoácidos. A descoberta das ribozimas complementa outras evidências que sugerem que o material genético original era o RNA.

É provável que as ribozimas autorreplicantes tenham surgido pela primeira vez entre 3,5 e 4 bilhões de anos atrás e tenham iniciado a evolução da vida na Terra. É provável que o início da vida tenha sido em um mundo de RNA, com as moléculas de RNA servindo como carreadoras de informação genética e como catalisadoras que controlavam as reações químicas necessárias para manter e perpetuar a vida. Esses RNA catalíticos podem ter adquirido a capacidade de sintetizar enzimas à base de proteínas, que são catalisadores mais eficientes. Com as enzimas dominando cada vez mais as funções catalíticas, é provável que o RNA tenha ficado relegado para o papel de armazenamento e transferência de informação. O DNA, com sua estabilidade química e replicação fidedigna, acabou substituindo o RNA como principal carreador de informação genética. Contudo, o RNA é produzido ou tem um papel vital em muitos processos biológicos, incluindo transcrição, replicação, processamento de RNA e tradução. A pesquisa nos últimos 15 anos também determinou que pequenas moléculas de RNA recém-descobertas têm participação fundamental em muitos processos biológicos, demonstrando que a vida hoje em dia ainda é muito parecida com um mundo de RNA. Essas pequenas moléculas de RNA serão discutidas com mais detalhes no Capítulo 14.

Conceitos

É provável que o início da vida tenha sido centrado no RNA, que serviu como material genético original e como catalisador biológico.

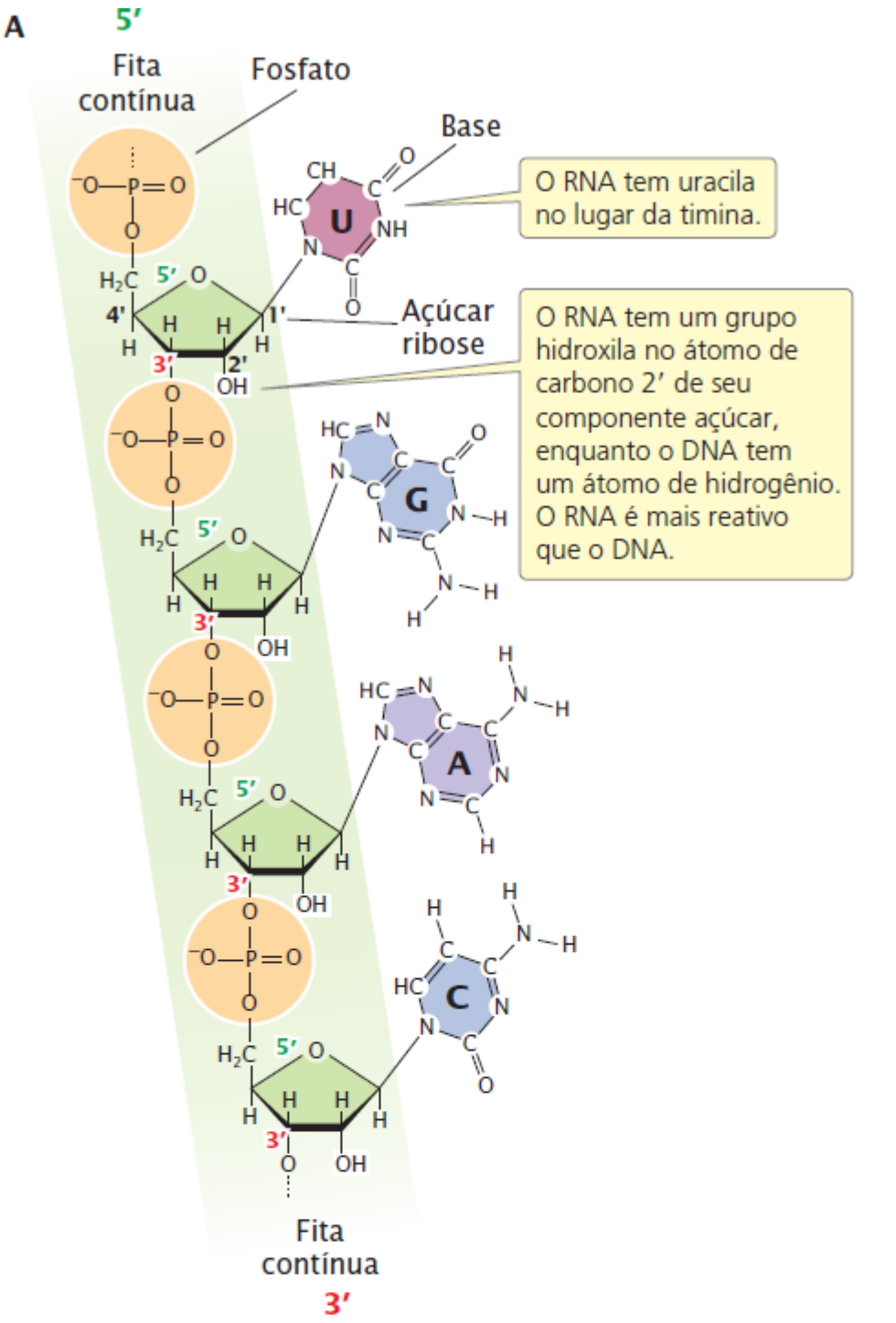
Estrutura do RNA

O RNA, como o DNA, é um polímero composto por nucleotídios unidos por ligações fosfodiéster (ler, no *Capítulo 10*, uma discussão sobre a estrutura do RNA). Entretanto, existem várias diferenças importantes nas estruturas do DNA e do RNA. Enquanto os nucleotídios do DNA têm açúcares desoxirribose, os nucleotídios do RNA têm ribose (**Figura 13.1 A**). Com um grupo hidroxila livre no átomo carbono 2′ do açúcar ribose, o RNA é rapidamente degradado em condições alcalinas. O açúcar desoxirribose do DNA não tem esse grupo hidroxila livre; por isso, o DNA é uma molécula mais estável. Outra diferença importante é que a timina, uma das duas pirimidinas encontradas no DNA, é substituída pela uracila no RNA.

Uma última diferença nas estruturas de DNA e RNA é que o RNA é composto por uma única fita de polinucleotídios (**Figura 13.1 B**), enquanto o DNA normalmente é composto por duas cadeias de polinucleotídios unidas por pontes de hidrogênio entre as bases complementares (apesar de alguns vírus terem genomas de RNA de fita dupla, como discutido no *Capítulo 9*). Embora o RNA tenha habitualmente fita simples, regiões complementares curtas em uma fita de nucleotídios podem parear e formar estruturas secundárias (ver Figura 13.1 B). Essas estruturas secundárias de RNA são chamadas de estruturas grampo-alça ou haste-alça. Quando as duas regiões em uma molécula simples de RNA pareiam, as fitas nessas regiões têm de ser antiparalelas, com paramento entre citosina e guanina e entre adenina e uracila (embora em alguns casos a guanina pareie com a uracila).

A formação de estruturas secundárias é importante na função do RNA. Elas são determinadas pela sequência de bases da fita de nucleotídios, então diferentes moléculas de RNA podem ter diferentes estruturas. Como sua estrutura determina sua função,

as moléculas de RNA têm o potencial para muitas funções diferentes. Com suas fitas complementares formando uma hélice, o DNA está muito mais restrito na faixa das estruturas secundárias que pode assumir, e por isso tem menos papéis funcionais na célula. O **Quadro 13.1** apresenta as semelhanças e diferenças entre as estruturas de DNA e RNA. **Resolva o Problema 14**


A
5′
Fita contínua
Fosfato
Base
O RNA tem uracila no lugar da timina.
Açúcar ribose
O RNA tem um grupo hidroxila no átomo de carbono 2′ de seu componente açúcar, enquanto o DNA tem um átomo de hidrogênio. O RNA é mais reativo que o DNA.
Fita contínua
3′

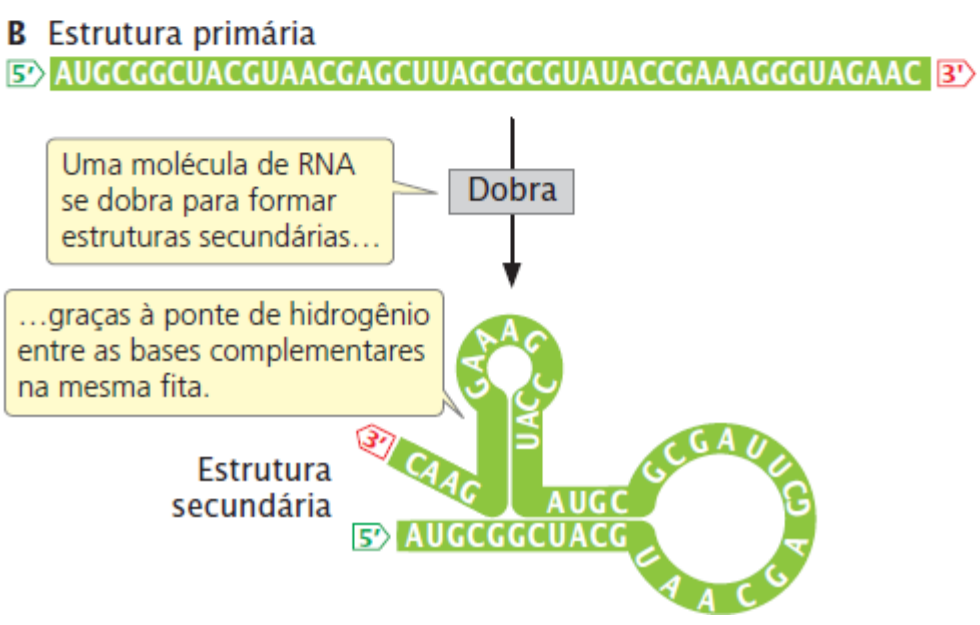

B Estrutura primária
5′ AUGCGGCUACGUAACGAGCUUAGCGCGUAUACCGAAAGGGUAGAAC 3′
Uma molécula de RNA se dobra para formar estruturas secundárias…
Dobra
…graças à ponte de hidrogênio entre as bases complementares na mesma fita.
Estrutura secundária
3′ CAAG AUGC GCGAUUCG
5′ AUGCGGCUACG UAACGA
GAAA UACC

Figura 13.1 **O RNA tem uma estrutura primária e uma estrutura secundária.**

Classes de RNA

As moléculas de RNA têm várias funções na célula. O **RNA ribossômico** (rRNA) e as subunidades ribossômicas de proteína compõem o ribossomo, o sítio de montagem das proteínas. Examinemos o ribossomo mais detalhadamente no Capítulo 14. O **RNA mensageiro** (mRNA) carreia as instruções de codificação para as cadeias de polipeptídios do DNA para um ribossomo. Após se prender ao ribossomo, uma molécula de mRNA especifica a sequência de aminoácidos em uma cadeia de polipeptídios e fornece um molde para unir os aminoácidos. As moléculas precursoras maiores, que são chamadas de **RNA pré-mensageiros** (pré-mRNAs), são os produtos imediatos da transcrição nas células eucarióticas. Os pré-mRNAs são muito modificados antes de se tornarem mRNA e de serem encontrados no núcleo para tradução em proteína. As células bacterianas não têm pré-mRNA; a transcrição ocorre simultaneamente com a tradução nessas células.

Quadro 13.1 Estruturas do DNA e RNA comparadas.

Característica	DNA	RNA
Composta por nucleotídios	Sim	Sim
Tipo de açúcar	Desoxirribose	Ribose
Grupo 2′-OH	Não	Sim
Bases	A, G, C, T	A, G, C, U
Nucleotídios unidos por ligações fosfodiéster	Sim	Sim
Fita dupla ou simples	Em geral, dupla	Em geral, simples
Estrutura secundária	Dupla-hélice	Muitos tipos
Estabilidade	Estável	Facilmente degradado

O **RNA transportador** (tRNA) serve como a ligação entre a sequência codificada de nucleotídios no mRNA e a sequência de aminoácidos de uma cadeia de polipeptídios. Cada tRNA se fixa a um tipo de aminoácido e o ajuda a ser incorporado em uma cadeia de polipeptídios (discutido no *Capítulo 15*).

Foram encontradas classes adicionais de moléculas de RNA nos núcleos das células eucarióticas. **RNAs nucleares pequenos** (snRNAs) se combinam com pequenas subunidades de proteínas para formar **ribonucleoproteínas nucleares pequenas** (snRNPs). Alguns snRNAs participam do processamento do RNA, convertendo o pré-mRNA em mRNA. **RNAs nucleolares pequenos** (snoRNAs) participam do processamento do rRNA.

Uma classe de moléculas de RNA muito pequena e abundante, chamada de **microRNAs** (miRNAs) e **RNAs de interferência pequenos** (siRNAs), encontra-se nas células eucarióticas e é responsável pelo RNA de interferência (RNAi), um processo no qual essas pequenas moléculas de RNA ajudam a acionar a degradação do mRNA ou inibem sua tradução em proteínas. O Capítulo 14 discutirá com mais detalhes o RNA de interferência. Uma pesquisa recente descobriu outra classe de pequenas moléculas de RNA, chamadas de **RNAs de interação com Piwi** (piRNAs, posteriormente denominadas proteínas Piwi, com as quais elas interagem). Encontradas nos testículos de mamíferos, essas moléculas de RNA se assemelham a miRNAs e siRNAs; acredita-se que participem na supressão da expressão dos elementos transponíveis (ver *Capítulo 18*) nas células reprodutoras. Recentemente, um sistema semelhante de interferência de RNA foi descoberto nos procariotos, nos quais pequenos **CRISPR RNAs** (crRNAs) auxiliam na destruição de moléculas de DNA estranhas. O **Quadro 13.2** resume algumas das diferentes classes de moléculas de RNA.

Quadro 13.2 Localização e funções de diferentes classes de moléculas de RNA.

Classe de RNA	Tipo de célula	Localização da função nas células eucarióticas*	Função
RNA ribossômico (rRNA)	Bacteriano e eucariótico	Citoplasma	Componentes estruturais e

			funcionais do ribossomo
RNA mensageiro (mRNA)	Bacteriano e eucariótico	Núcleo e citoplasma	Carreia o código genético para proteínas
RNA transportador (tRNA)	Bacteriano e eucariótico	Citoplasma	Ajuda a incorporar os aminoácidos na cadeia de polipeptídios
RNA nuclear pequeno (snRNA)	Eucariótico	Núcleo	Processamento do pré-mRNA
RNA nucleolar pequeno (snoRNA)	Eucariótico	Núcleo	Processamento e montagem do rRNA
MicroRNA (miRNA)	Eucariótico	Núcleo e citoplasma	Inibe a tradução do mRNA
RNA de interferência pequeno (siRNA)	Eucariótico	Núcleo e citoplasma	Aciona a degradação de outras moléculas de RNA
RNA de interação com Piwi (piRNA)	Eucariótico	Núcleo e citoplasma	Suprime a transcrição dos elementos transponíveis nas células reprodutoras
CRISPR RNA (crRNA)	Procariótico	–	Auxilia na destruição de DNA estranho

*Todos os RNAs eucarióticos são sintetizados no núcleo.

Conceitos

O RNA é diferente do DNA, pois apresenta um grupo hidroxila no átomo carbono 2′ do seu açúcar, tem uracila em vez da timina e, em geral, é uma fita simples. Existem várias classes de RNA nas células bacterianas e eucarióticas.

Checagem dos conceitos 1

Qual classe de RNA está corretamente combinada com sua função?

a. RNA nuclear pequeno (snRNA): processa rRNA.
b. RNA transportador (tRNA): fixa-se a um aminoácido.
c. MicroRNA (miRNA): carrega a informação genética para a sequência de aminoácidos de uma proteína.
d. RNA ribossômico (rRNA): faz a interferência do RNA.

13.2 Transcrição é a síntese de uma molécula de RNA a partir de um molde de DNA

Todos os RNAs celulares são sintetizados a partir de moldes de DNA por meio do processo de transcrição (**Figura 13.2**). A transcrição tem muitas semelhanças com o processo de replicação, mas a diferença fundamental está relacionada com o tamanho do molde usado. Na replicação, todos os nucleotídios do molde do DNA são copiados, mas, na transcrição, apenas pequenas partes da molécula de DNA – em geral um único gene, ou, no máximo, alguns genes – são transcritos em RNA. Como nem todos os produtos dos genes são necessários ao mesmo tempo ou na mesma célula, a transcrição constante de todos os genes de uma célula seria muito ineficiente. Além disso, boa parte do DNA não codifica um produto funcional e a transcrição de tais sequências não faria sentido. A transcrição é, de fato, um processo muito seletivo: os genes individuais são transcritos apenas se seus produtos forem necessários. Entretanto, essa seletividade impõe um problema fundamental na célula: como identificar os genes individuais e transcrevê-los no momento e local corretos.

Como a replicação, a transcrição precisa de três componentes principais:

1. Um molde de DNA.
2. Matérias-primas (trifosfato de ribonucleotídio) necessárias para construir uma molécula de RNA.
3. O aparato de transcrição, com as proteínas necessárias para catalisar a síntese de RNA.

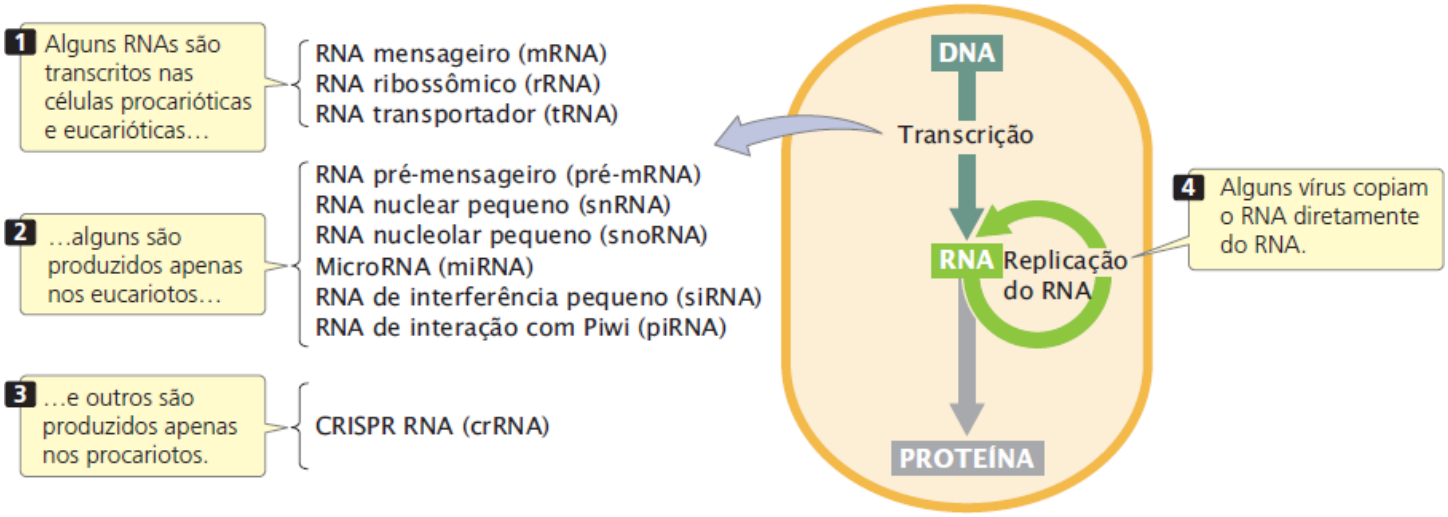


Figura 13.2 Todos os tipos celulares de RNA são transcritos a partir do DNA.

Molde

Em 1970, Oscar Miller Jr., Barbara Hamkalo e Charles Thomas usaram microscopia eletrônica para examinar o material celular e demonstrar que o RNA é transcrito a partir de um molde de DNA. Eles viram estruturas semelhantes à árvore de Natal dentro da célula: finas fibras centrais (o tronco da árvore) às quais estavam presos fios (os ramos) com grânulos (**Figura 13.3 A**). A adição da desoxirribonuclease (uma enzima que degrada o DNA) fez com que as fibras centrais desaparecessem, indicando que os "troncos da árvore" eram moléculas de DNA. A ribonuclease (uma enzima que degrada RNA) removeu os fios granulosos, indicando que os ramos eram RNA. Sua conclusão foi que cada "árvore de Natal" representava um gene sendo submetido à transcrição (**Figura 13.3 B**). A transcrição de cada gene inicia-se no topo da árvore; então, pouco do DNA foi transcrito e os ramos de RNA eram curtos. À medida que o aparato de transcrição desce na árvore, transcrevendo mais o molde, as moléculas de RNA se prolongam, produzindo grandes ramificações no final da árvore.

A fita transcrita. O molde para a síntese do RNA, assim como para a síntese de DNA, é uma fita simples da dupla-hélice do DNA. Diferente da replicação, entretanto, a transcrição de um gene ocorre em apenas uma das duas fitas de nucleotídios do DNA (**Figura 13.4**). A fita de nucleotídios usada para a transcrição é chamada de **fita molde**. A outra, chamada de **fita não molde**, não é transcrita. Assim, em um gene, apenas uma das fitas de nucleotídios é transcrita em RNA (existem algumas exceções para esta regra).

Durante a transcrição, uma molécula de RNA complementar e antiparalela à fita molde de DNA é sintetizada (ver Figura 13.4). O RNA transcrito tem a mesma polaridade e sequência de bases que a fita não molde, com a exceção de que o RNA tem U em vez de T. Na maioria dos organismos, cada gene é transcrito a partir de uma fita simples, mas diferentes genes podem ser transcritos a partir de diferentes fitas, como apresentado na **Figura 13.5**. Resolva o Problema 15

Conceitos

Em um único gene, apenas uma das duas fitas de DNA, a fita molde, é transcrita em RNA.

Checagem dos conceitos 2

Qual é a diferença entre a fita molde e a fita não molde?

Unidade de transcrição. Uma **unidade de transcrição** é um segmento de DNA que codifica uma molécula de RNA e as sequências necessárias para sua transcrição. Como o complexo de enzimas e proteínas que faz a transcrição – o aparato de transcrição – identifica uma unidade de transcrição? Como ele sabe qual fita de DNA a ser lida e onde começar a parar? Essa informação é codificada pela sequência de DNA.

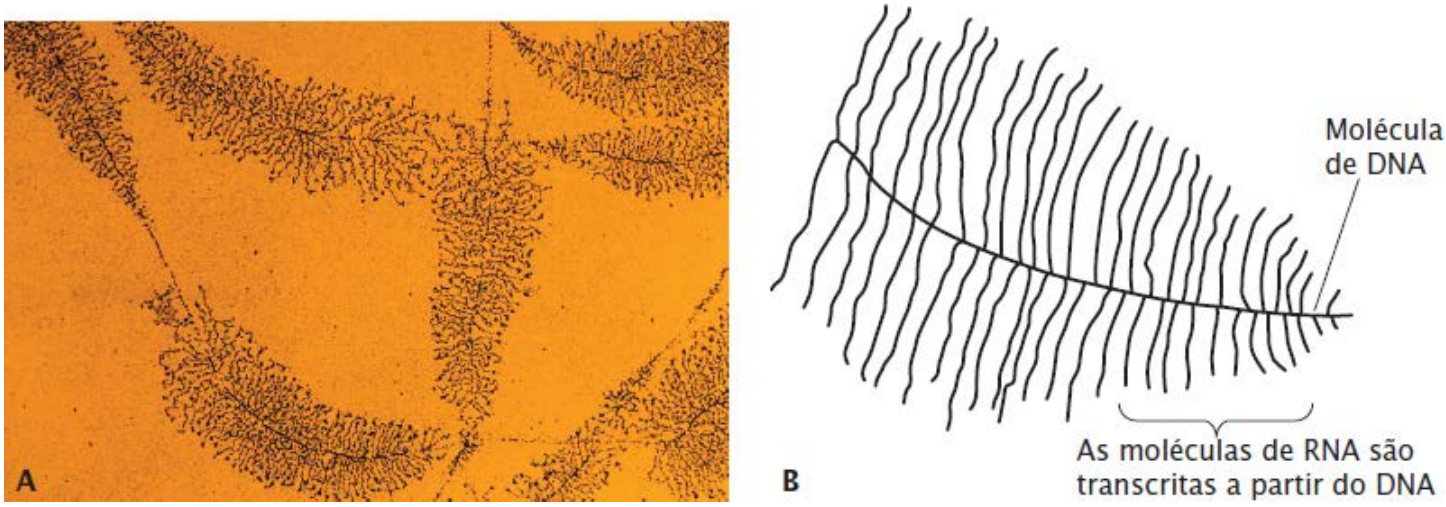


Figura 13.3 À microscopia eletrônica, as moléculas de DNA submetidas à transcrição exibem estruturas semelhantes à árvore de Natal. A. Micrografia eletrônica de estruturas semelhantes a árvores de Natal. **B.** O tronco de cada "árvore de Natal" (uma unidade de transcrição) representa uma molécula de DNA; as ramificações da árvore são moléculas de RNA que foram transcritas do DNA. À medida que o aparato de transcrição desce na árvore, transcrevendo mais do molde, as moléculas de RNA se tornam cada vez mais compridas (*Parte A*: Dr. Thomas Broker/Phototake.)

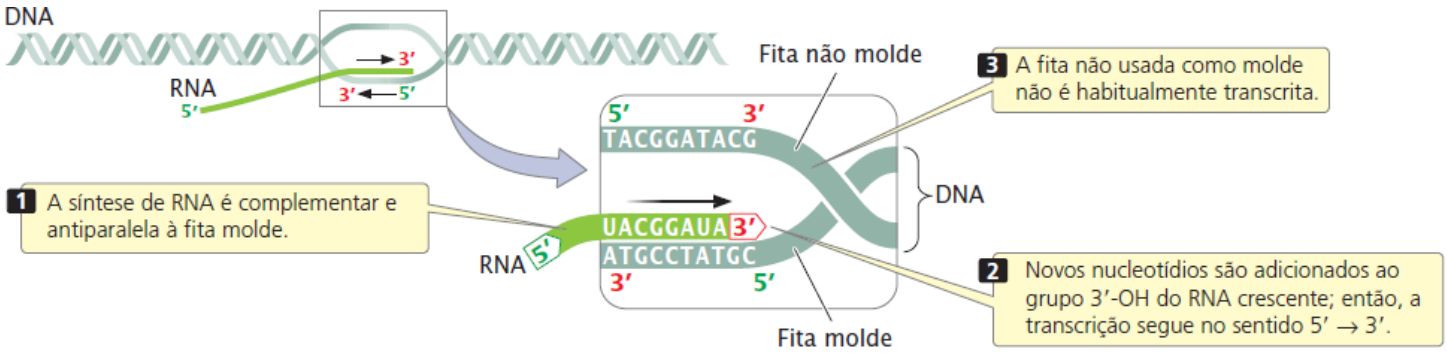


Figura 13.4 As moléculas de RNA sintetizadas são complementares e antiparalelas a uma das duas fitas de nucleotídios do DNA, a fita molde.

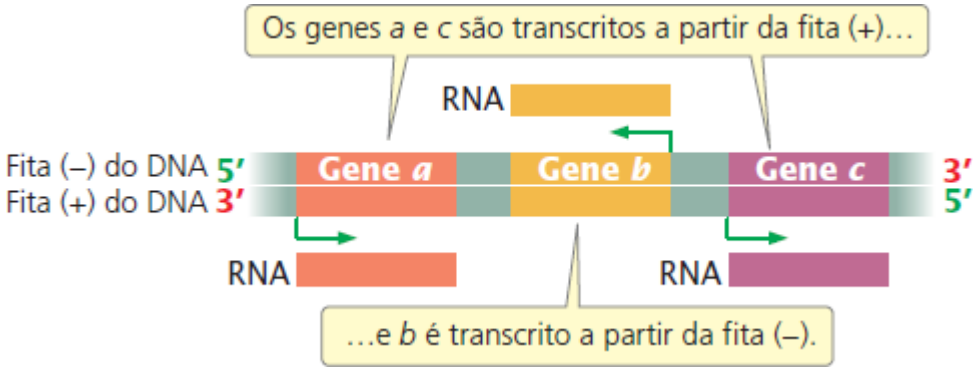


Figura 13.5 O RNA é transcrito a partir de uma fita de DNA. Na maioria dos organismos, cada gene é transcrito a partir de uma fita simples de DNA, mas diferentes genes podem ser transcritos a partir de cada fita de DNA.

Três regiões críticas estão incluídas em uma unidade de transcrição: um promotor, uma sequência codificadora de RNA e um terminador (**Figura 13.6**). O **promotor** é uma sequência de DNA que o aparato de transcrição reconhece e se liga. Ele indica qual das duas fitas de DNA deve ser lida como molde e qual será o sentido da transcrição. O promotor também determina o sítio de início da transcrição, o primeiro nucleotídio que será transcrito em RNA. Em muitas unidades de transcrição, o promotor está próximo ao sítio de início de transcrição, mas ele mesmo não é transcrito.

A segunda região crítica da unidade de transcrição é a **região codificadora de RNA**, uma sequência de nucleotídios de DNA que é copiada em uma molécula de RNA. O terceiro componente da unidade de transcrição é o **terminador**, uma sequência de nucleotídios que sinaliza onde a transcrição deve terminar. Os terminadores são parte da sequência codificadora de RNA; a transcrição para somente após o terminador ter sido copiado em RNA.

Os biólogos moleculares usam os termos *upstream* e *downstream* para se referir ao sentido da transcrição e à localização das

sequências de nucleotídios que circundam a sequência codificadora de RNA. O aparato de transcrição se desloca *downstream* durante a transcrição: ele se liga ao promotor (que em geral está *upstream* do sítio de início) e se move para o terminador (que está *downstream* do sítio de início).

Quando as sequências de DNA estiverem transcritas, com frequência apenas a sequência de uma das duas fitas é registrada. Os biólogos moleculares escrevem a sequência da fita não molde, porque ela será a mesma que a sequência do RNA transcrito a partir do molde (com exceção de que U no RNA substitui T no DNA). Por convenção, a sequência da fita não molde é escrita com a extremidade 5′ à esquerda e a extremidade 3′ à direita. O primeiro nucleotídio transcrito (o sítio de início de transcrição) é numerado +1; os nucleotídios *downstream* do sítio de início são indicados com números positivos, e os nucleotídios *upstream* do sítio de início são indicados com números negativos. Então, o nucleotídio +34 estaria 34 nucleotídios *downstream* do sítio de início, enquanto o nucleotídio −75 estaria 75 nucleotídios *upstream* do sítio de início. Não existe nucleotídio com o número 0.

Conceitos

A unidade de transcrição é um segmento de DNA que codifica uma molécula de RNA e as sequências necessárias para sua transcrição. Cada unidade de transcrição inclui um promotor, uma região codificadora de RNA e um terminador.

Checagem dos conceitos 3

Qual das frases a seguir *não* descreve a função do promotor?

a. Serve como sequência à qual o aparato de transcrição se liga.
b. Determina o primeiro nucleotídio que é transcrito em RNA.
c. Determina qual fita de DNA é o molde.
d. Sinais onde a transcrição cessa.

Substrato para transcrição

O RNA é sintetizado a partir de **trifosfatos de ribonucleosídios** (rNTPs, **Figura 13.7**). Na síntese, novos nucleotídios são unidos, um de cada vez, ao grupo 3′-OH da molécula de RNA crescente. Dois grupos fosfato são rompidos a partir do trifosfato de nucleosídio de chegada; o grupo fosfato remanescente participa em uma ligação fosfodiéster que liga o nucleotídio à molécula de RNA crescente. A reação química completa para a adição de cada nucleotídio é:

$$RNA_n + rNTP \rightarrow RNA_{n+1} + PP_i$$

em que PP_i representa pirofosfato. Os nucleotídios são sempre adicionados à extremidade 3′ da molécula de RNA, e o sentido da transcrição é, portanto, 5′ → 3′ (**Figura 13.8**), o mesmo que a síntese de DNA durante a replicação. A síntese de RNA é complementar e antiparalela a uma das fitas de DNA (a fita molde). Diferente da síntese do DNA, a síntese do RNA não precisa de um *primer*.

Conceitos

O RNA é sintetizado a partir de trifosfatos de ribonucleosídios. A transcrição é 5′ → 3′: cada novo nucleotídio é unido ao grupo 3′-OH do último nucleotídio adicionado à molécula de RNA crescente.

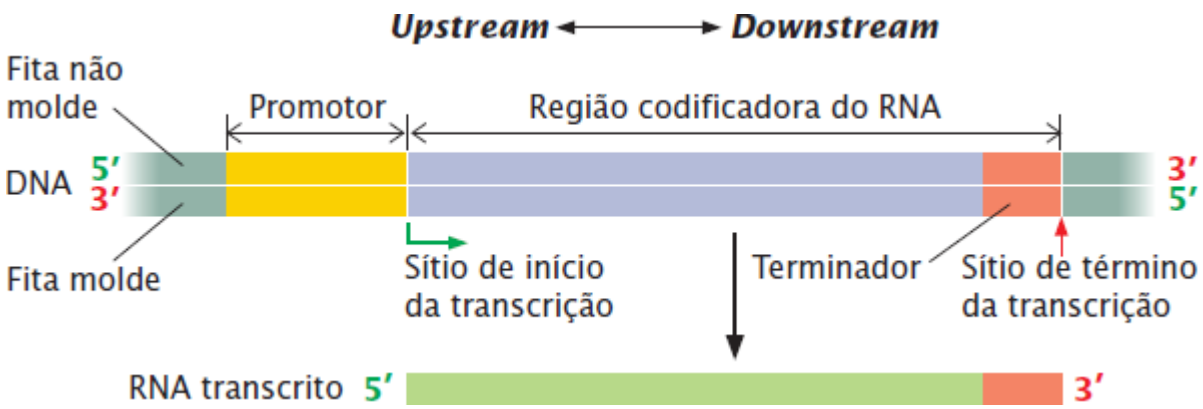


Figura 13.6 **Uma unidade de transcrição inclui um promotor, uma região codificadora de RNA e um terminador.**

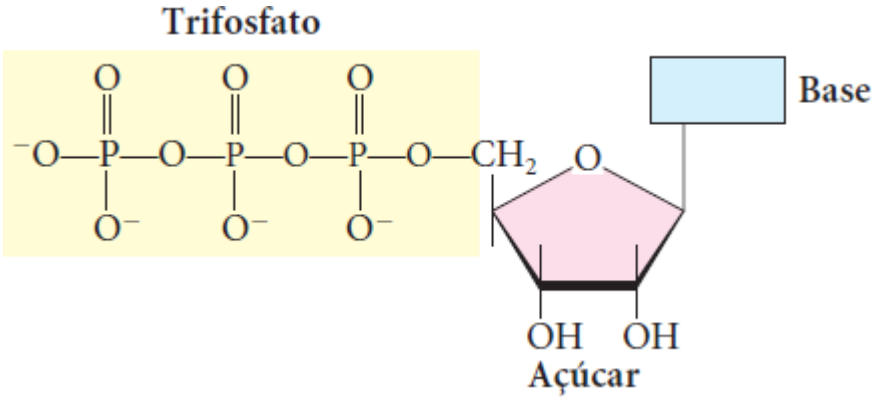


Figura 13.7 **Os trifosfatos de ribonucleosídios são substratos usados na síntese de RNA.**

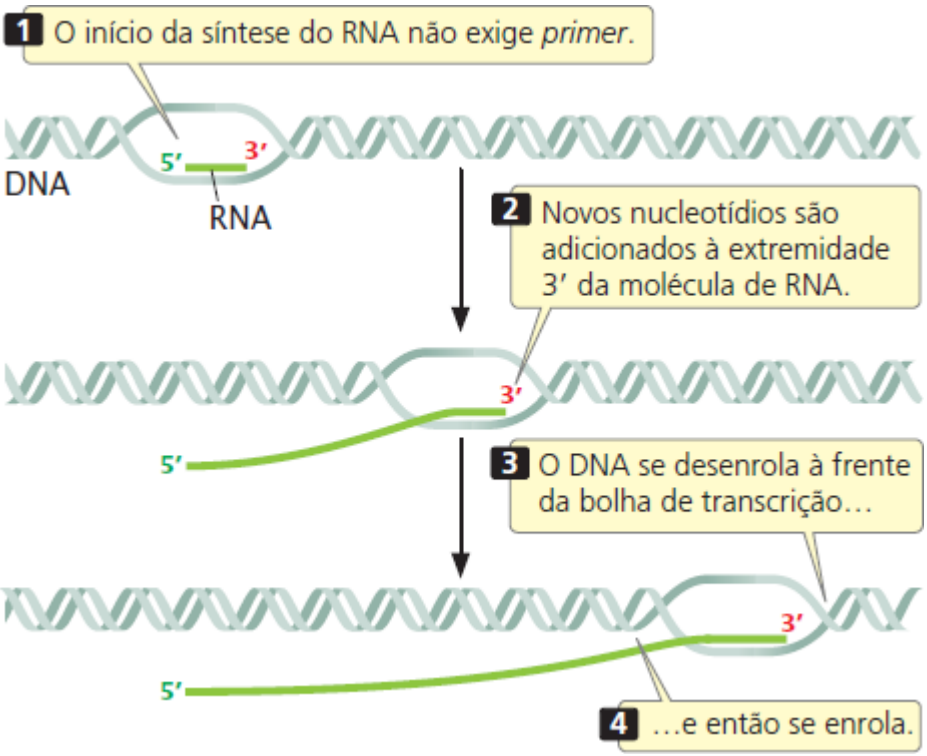


Figura 13.8 **Na transcrição, novos nucleotídios são sempre adicionados à extremidade 3′ da molécula de RNA.**

Aparato de transcrição

Lembre-se de que a replicação do DNA precisa de várias enzimas e proteínas diferentes. Embora a transcrição possa parecer bem diferente inicialmente porque uma única enzima, a **RNA polimerase**, realiza todas as etapas necessárias da transcrição, um exame mais meticuloso revela que os processos são semelhantes. A ação da RNA polimerase é potencializada por várias proteínas acessórias que se unem e deixam a polimerase em diferentes estágios do processo. Cada proteína acessória é responsável por fornecer ou regular uma função especial. Assim, a transcrição, como a replicação, requer um conjunto de proteínas.

RNA polimerase bacteriana. As células bacterianas têm tipicamente apenas um tipo de RNA polimerase, que catalisa a síntese de todas as classes de RNA bacteriano: mRNA, tRNA e rRNA. A RNA polimerase bacteriana é uma enzima grande, multimérica (o que significa que tem várias cadeias de polipeptídios).

Existem cinco subunidades (cadeias individuais de polipeptídios) na parte central da maioria das RNA polimerases bacterianas que compõem o **cerne da enzima**: duas cópias de uma subunidade chamada alfa (**α**), e cópias únicas das subunidades beta (β), beta linha (β′) e ômega (**ω**) (**Figura 13.9**). A subunidade ω não é essencial para a transcrição, mas ajuda a estabilizar a enzima. O cerne da enzima catalisa o prolongamento da molécula de RNA pela adição de nucleotídios de RNA. Outras subunidades funcionais unem e deixam o cerne da enzima em estágios específicos do processo de transcrição. O **fator sigma** (**σ**) controla a ligação da RNA polimerase ao promotor. Sem o sigma, a RNA polimerase inicia a transcrição em um ponto aleatório ao longo do DNA. Após o sigma se associar com o cerne da enzima (formando a **holoenzima**), a RNA

polimerase se liga de forma estável apenas à região do promotor e inicia a transcrição no sítio de início correto. O sigma é necessário apenas para a ligação ao promotor e para o início; quando alguns nucleotídios de RNA são unidos, em geral, ele se solta do cerne da enzima. Muitas bactérias apresentam múltiplos tipos de fatores sigma; cada tipo de sigma inicia a ligação à RNA polimerase em um conjunto particular de promotores.

As rifamicinas são um grupo de antibióticos que exterminam as bactérias ao inibir a RNA polimerase. Esses antibióticos são muito usados para tratar tuberculose, uma doença que mata quase 2 milhões de pessoas pelo mundo todo ano. As estruturas das RNA polimerases bacterianas e eucarióticas são diferentes o suficiente para que as rifamicinas inibam as RNA polimerases bacterianas sem interferir com as RNA polimerases eucarióticas. Pesquisa recente demonstrou que várias rifamicinas inibem a RNA polimerase ao se ligar à parte da enzima que se prende ao DNA, impedindo que a enzima interaja com o promotor no DNA.

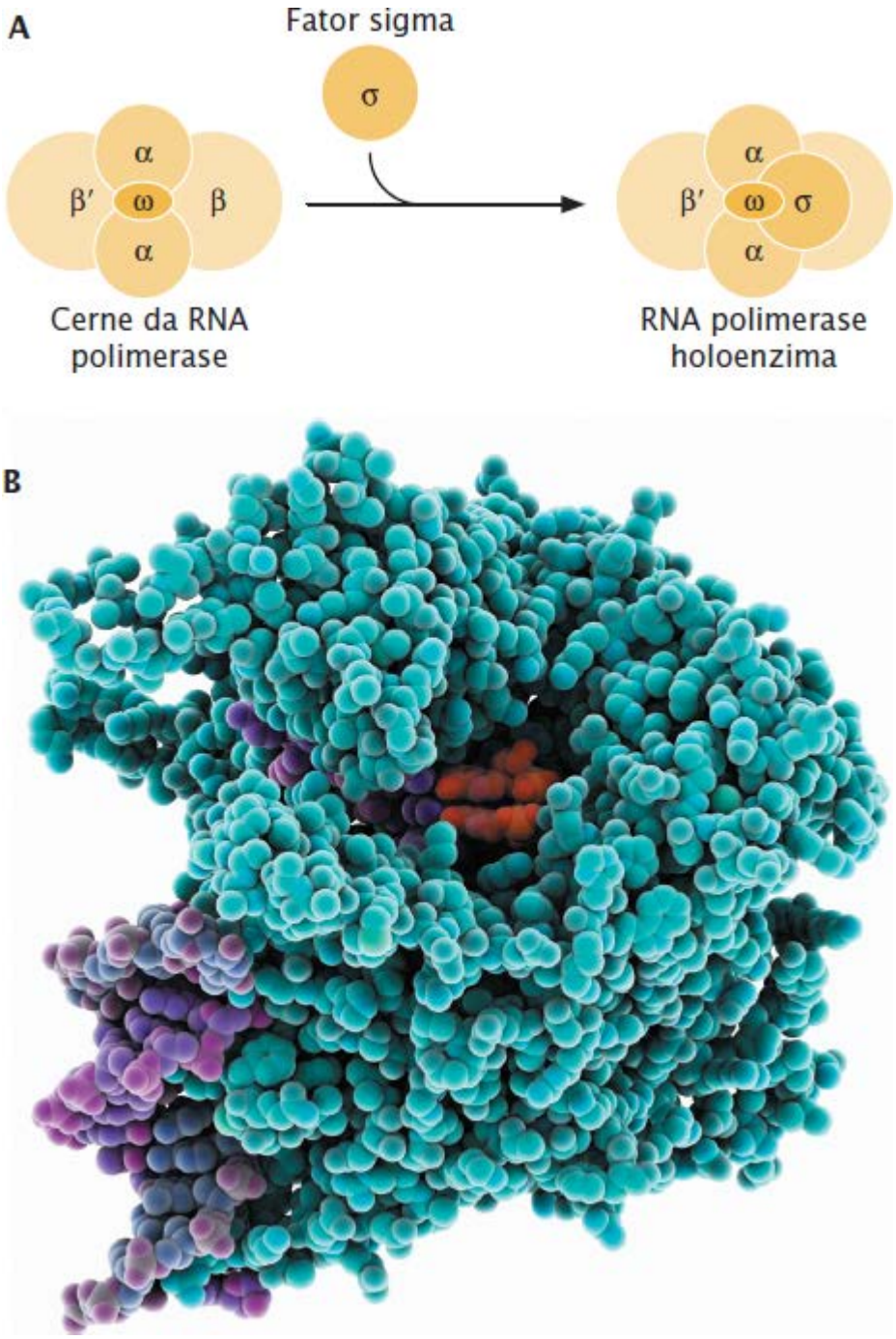


Figura 13.9 Na RNA polimerase bacteriana, o cerne da enzima é composto por cinco subunidades: duas cópias de alfa (*α*), uma única cópia de beta (*β*), uma única cópia de beta *prime* (*β′*) e uma única cópia de ômega (*ω*). O cerne da enzima catalisa o prolongamento da molécula de RNA pela adição de nucleotídios de RNA. **A.** O fator sigma (*σ*) se une ao cerne da enzima para formar a holoenzima, que é capaz de se ligar a um promotor e iniciar a transcrição. **B.** O modelo molecular mostra a RNA polimerase (*azul*), ligando-se ao DNA (*roxo*) e sintetizando mRNA (*vermelho*). (*Parte B*: Laguna Design/Science Photo Library.)

RNA polimerases eucarióticas. A maioria das células eucarióticas tem três tipos diferentes de RNA polimerase, cada uma responsável por transcrever uma classe diferente de RNA: a **RNA polimerase I** transcreve rRNA; a **RNA polimerase II**

transcreve pré-mRNAs; snoRNAs, alguns miRNAs e alguns snRNAs; e a **RNA polimerase III** transcreve outras moléculas de RNA pequenas – especificamente tRNAs, rRNA pequeno, alguns miRNAs e alguns snRNAs (**Quadro 13.3**). As RNA polimerases I, II e III são encontradas em todos os eucariotos. Foram encontradas duas RNA polimerases adicionais nas plantas, **RNA polimerase IV** e **RNA polimerase V**. As RNA polimerases IV e V transcrevem RNAs que atuam na metilação do DNA e na estrutura da cromatina.

Todas as polimerases eucarióticas são enzimas grandes, multiméricas, com mais de uma dúzia de subunidades. Algumas subunidades são comuns a todas as RNA polimerases, enquanto outras são limitadas a uma das polimerases. Como em todas as células bacterianas, várias proteínas acessórias se ligam ao cerne da enzima e afetam sua função.

Conceitos

As células bacterianas têm um tipo único de RNA polimerase, composta por um cerne e outras subunidades que participam em vários estágios da transcrição. As células eucarióticas têm vários tipos diferentes de RNA polimerase que transcrevem diferentes tipos de moléculas de RNA.

Checagem dos conceitos 4

Qual é a função do fator sigma?

13.3 A transcrição bacteriana consiste em iniciação, alongamento e término

Agora que analisamos alguns dos principais componentes da transcrição, estamos prontos para olhar com detalhes o processo. A transcrição pode ser dividida em três estágios:

1. Iniciação, na qual o aparato da transcrição se encaixa no promotor e inicia a síntese de RNA.
2. Alongamento, no qual o DNA é inserido na RNA polimerase, a polimerase abre o DNA e adiciona novos nucleotídios, um por vez, à extremidade 3′ da fita de RNA crescente.
3. Término, a identificação da extremidade da unidade de transcrição e a separação da molécula de RNA do molde de DNA.

Quadro 13.3 RNA polimerases eucarióticas.

Tipo	Encontrada em	Transcreve
RNA polimerase I	Todos os eucariotos	rRNAs grandes
RNA polimerase II	Todos os eucariotos	Pré-mRNA, alguns snRNAs, snoRNAs, alguns miRNAs
RNA polimerase III	Todos os eucariotos	tRNAs, rRNAs pequenos, alguns snRNAs, alguns miRNAs
RNA polimerase IV	Plantas	Alguns siRNAs
RNA polimerase V	Plantas	As moléculas de RNA participam da formação da heterocromatina

Primeiro examinaremos cada uma dessas etapas nas células bacterianas, onde o processo é mais bem compreendido; então, analisaremos a transcrição eucariótica e de archaea.

Iniciação

A iniciação abrange todas as etapas necessárias para começar a síntese do RNA, incluindo (1) identificação do promotor, (2) formação da bolha de transcrição, (3) criação das primeiras ligações entre rNTPs e (4) saída do aparato de transcrição do promotor.

A iniciação da transcrição requer que o aparato de transcrição identifique e se ligue ao promotor. Nessa etapa, a seletividade da transcrição é reforçada; a ligação da RNA polimerase ao promotor determina quais partes do molde de DNA devem ser transcritos e com que frequência. Diferentes genes são transcritos com diferentes frequências, e a ligação ao promotor é responsável principalmente por determinar a frequência de transcrição para um gene particular. Os promotores também têm diversas afinidades pela RNA polimerase. Mesmo dentro de um único promotor, a afinidade pode variar com o tempo, dependendo da interação do promotor com a RNA polimerase e de vários outros fatores.

Promotores bacterianos. A informação essencial para a unidade de transcrição – onde se inicia a transcrição, qual fita deve ser lida e em qual sentido da RNA polimerase se move – está incrustada na sequência de nucleotídios do promotor. Os promotores são sequências de DNA identificadas pelo aparato de transcrição e necessárias à transcrição. Nas células bacterianas, os promotores estão adjacentes a uma sequência codificadora de RNA.

Um exame de vários promotores na *E. coli* e em outras bactérias revelou uma característica geral: embora a maioria dos nucleotídios dentro dos promotores varie na sequência, os segmentos curtos de nucleotídios são comuns em muitos. Além disso, a distância e a localização desses nucleotídios relativos ao sítio de início de transcrição são semelhantes na maioria dos promotores. Esses trechos curtos dos nucleotídios comuns são chamados de **sequências consenso**; "sequência consenso" refere-se a sequências com grande semelhança ou consenso (**Figura 13.10**). O consenso em um conjunto de nucleotídios, em geral, indica que a sequência está associada a uma função importante. Resolva o Problema 21

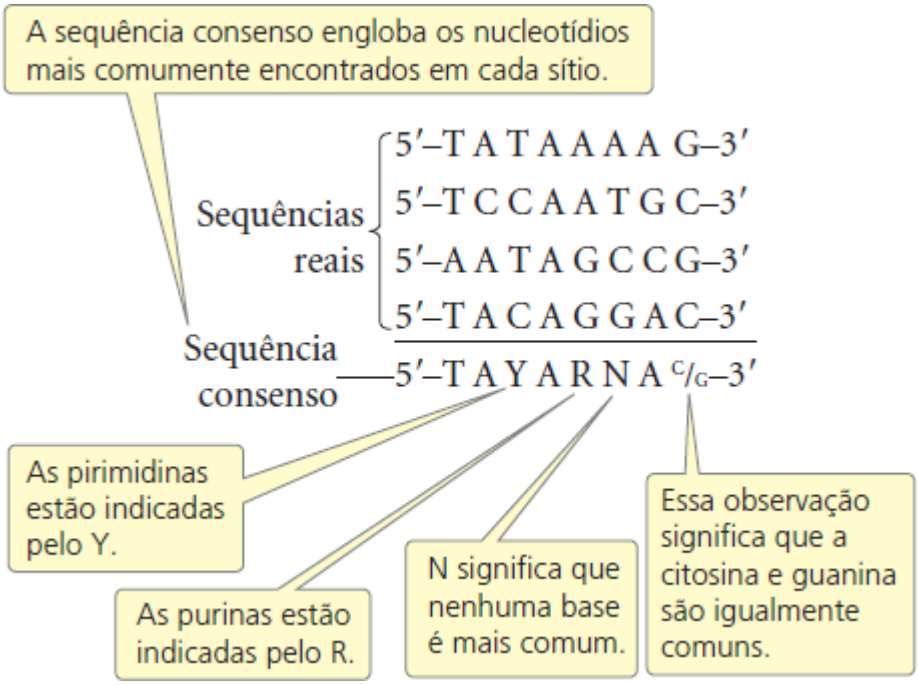


Figura 13.10 Uma sequência consenso é composta por bases mais comumente encontradas em cada posição em um grupo de sequências próximas.

A sequência consenso mais comum, encontrada na maioria dos promotores bacterianos, está centralizada cerca de 10 pb *upstream* do sítio de início. Chamado de **sequência consenso −10** ou, às vezes, a Pribnow box, sua sequência consenso é

5′–T A T A A T–3′
3′–A T A T T A–5′

e é frequentemente escrita apenas como TATAAT (**Figura 13.11**). Lembre-se de que TATAAT é apenas a sequência *consenso* – representando a maioria dos nucleotídios comumente encontrados em cada uma dessas posições. Na maior parte dos promotores procarióticos, a sequência real não é TATAAT.

Outra sequência consenso comum à maioria dos promotores bacterianos é TTGACA, que fica a aproximadamente 35 nucleotídios *upstream* do sítio de início e é chamada de **sequência consenso −35** (ver Figura 13.11). Os nucleotídios de cada lado das sequências consenso −10 e −35 e os que estão entre elas variam muito de promotor para promotor, sugerindo que esses nucleotídios não são muito importantes na identificação do promotor.

A função dessas sequências consenso nos promotores bacterianos foi estudada ao induzir mutações em várias posições dentro dessas sequências e observando o efeito das mudanças na transcrição. Os resultados desses estudos revelam que a maioria das substituições de base dentro das sequências consenso −10 e −35 reduz a taxa de transcrição; essas substituições são chamadas

de *mutações redutoras* (*down mutations*) porque reduzem a taxa de transcrição. Em alguns casos, uma mudança particular em uma sequência consenso aumenta a taxa de transcrição; tal mudança é chamada de *mutação ativadora* (*up mutation*).

Como mencionado, o fator sigma se associa com o cerne da enzima (**Figura 13.12 A**) para formar uma holoenzima, que se liga às sequências consenso −35 e −10 no promotor de DNA (**Figura 13.12 B**). Embora se ligue apenas aos nucleotídios das sequências consenso, a enzima se estende de −50 a +20 quando se liga ao promotor. Inicialmente, a holoenzima faz uma ligação fraca com o promotor, mas então sofre mudança na estrutura que possibilita uma ligação com mais intensidade e o desenrolar da fita dupla do DNA (**Figura 13.12 C**). O desenrolamento começa dentro da sequência consenso −10 e se estende *downstream* por cerca de 14 nucleotídios, incluindo o sítio de início (a partir dos nucleotídios −12 a +2).

Alguns promotores bacterianos têm uma terceira sequência consenso que também participa do início da transcrição. Chamada de **elemento *upstream***, essa sequência tem vários pares A-T e é encontrada cerca de −40 a −60. Várias proteínas podem ligar-se a sequências dentro e próximas do promotor; algumas estimulam a taxa de transcrição, e outras a reprimem. Estudaremos essas proteínas, que regulam a expressão gênica, no Capítulo 16. Resolva o Problema 24

Conceitos

Um promotor é uma sequência de DNA adjacente ao gene e necessário para a transcrição. Os promotores têm sequências consenso curtas importantes para o início da transcrição.

Checagem dos conceitos 5

O que se liga à sequência consenso −10 encontrada na maioria dos promotores bacterianos?

a. A holoenzima (cerne da enzima + sigma).
b. O fator sigma sozinho.
c. O cerne da enzima sozinho.
d. mRNA.

Síntese inicial do RNA. Após a holoenzima se prender ao promotor, a RNA polimerase é posicionada sobre o sítio de início para transcrição (na posição +1) e desenrola o DNA para produzir um molde de fita simples. A orientação e a distância das sequências consenso em uma fita de DNA determinam qual fita será o molde para a transcrição e, portanto, determina o sentido da transcrição.

A posição do sítio de início é determinada não apenas pelas sequências localizadas nessa região, mas também pela localização das sequências consenso, que posicionam a RNA polimerase de modo que o sítio ativo da enzima esteja alinhado para iniciar a transcrição em +1. Se as sequências consenso são artificialmente removidas *upstream* ou *downstream*, a localização do ponto de início da transcrição se altera.

Para iniciar a síntese de uma molécula de RNA, a RNA polimerase pareia a base em um trifosfato de ribonucleosídio com sua base complementar no sítio de início na fita molde de DNA (**Figura 13.12 D**). Não é necessário nenhum *primer* para iniciar a síntese na extremidade 5′ da molécula de RNA. Dois dos três grupos de fosfato são rompidos a partir do trifosfato de ribonucleosídio à medida que o nucleotídio é adicionado à extremidade 3′ da molécula de RNA crescente. Entretanto, como a extremidade 5′ do primeiro trifosfato de ribonucleosídio não participa da formação de uma ligação fosfodiéster, todos os três grupos fosfatos são preservados. Uma molécula de RNA, portanto, tem pelo menos inicialmente, três grupos fosfatos na sua extremidade 5′ (**Figura 13.12 E**).

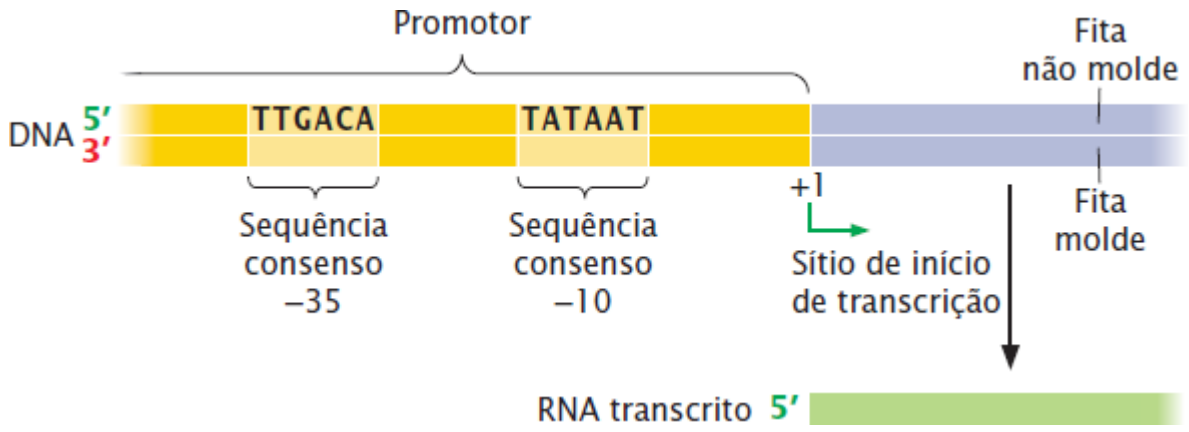


Figura 13.11 **Nos promotores bacterianos, as sequências consenso estão *upstream* do sítio de início, aproximadamente nas**

posições −10 e −35.

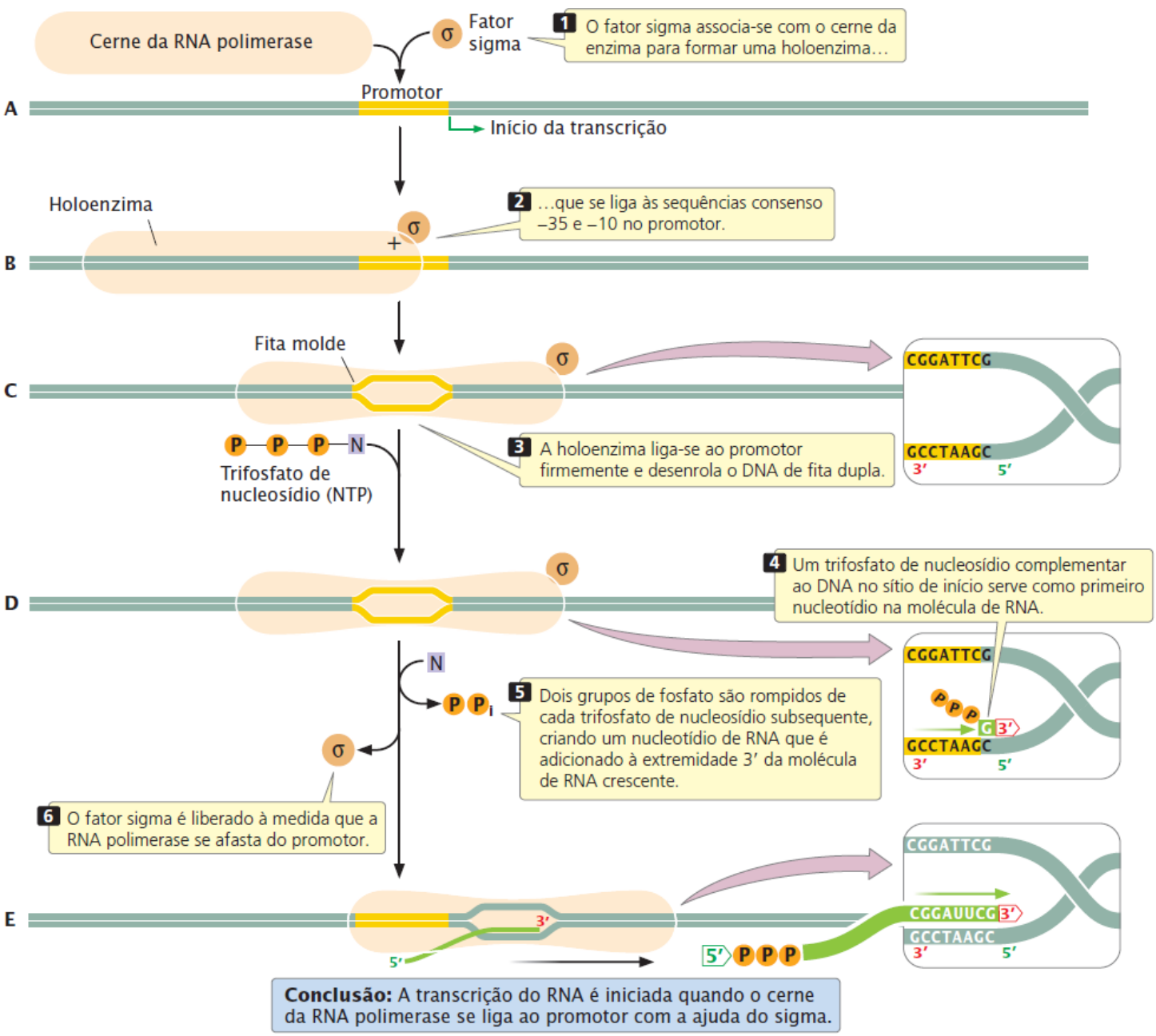


Figura 13.12 A transcrição nas bactérias é feita pela RNA polimerase, que precisa ligar-se ao fator sigma para iniciar a transcrição.

Durante o início da transcrição, a RNA polimerase gera e libera repetidamente transcrições curtas, de 2 a 6 nucleotídios de comprimento, enquanto ainda está presa ao promotor. Esse processo, chamado **início abortado**, ocorre tanto nos procariotos quanto nos eucariotos. Após várias tentativas abortadas, a polimerase sintetiza uma molécula de RNA com 9 a 12 nucleotídios de comprimento, o que torna possível a transição da RNA polimerase para o estágio de alongamento.

Alongamento

Após a iniciação da transcrição, a RNA polimerase sofre uma mudança na conformação (formato) e não é mais capaz de se ligar a sequências consenso no promotor. Essa mudança torna possível que a polimerase saia do promotor e comece a transcrição *downstream*. A subunidade sigma é liberada após o início, embora algumas populações de RNA polimerase possam reter o sigma durante todo o alongamento.

À medida que ela se move *downstream* ao longo do molde, a RNA polimerase desenrola o DNA na borda líder (*downstream*) da bolha de transcrição, unindo nucleotídios à molécula de RNA de acordo com a sequência no molde, e enrola o DNA na fronteira *upstream* da bolha. Cerca de 40 nucleotídios por segundo são adicionados nas células bacterianas a 37°C. Essa taxa de síntese de RNA é muito menor do que a da síntese de DNA, que é de 1.000 a 2.000 nucleotídios por segundo nesse tipo de células.

Bolha de transcrição. A transcrição ocorre dentro de um curto segmento de 18 nucleotídios de DNA desenrolado – a bolha de transcrição. Dentro dessa região, o RNA é sintetizado de forma contínua, com o DNA de fita simples usado como molde. Cerca de 8 nucleotídios de RNA recém-sintetizado são pareados com os nucleotídios do molde de DNA em qualquer momento. À medida que o aparato de transcrição desce o molde de DNA, ele gera um superdesenrolamento positivo à frente da bolha de transcrição e um superenrolamento negativo atrás dela. É provável que as enzimas topoisomerases aliviem a tensão associada ao desenrolamento e enrolamento do DNA na transcrição, como fazem na replicação do DNA.

Pausa transcricional. Várias características do RNA ou do DNA, como estruturas secundárias, sequências específicas ou a existência de nucleossomos, fazem com que a RNA polimerase pause o estágio de alongamento da transcrição. As pausas são causadas pelo recuo – quando a RNA polimerase desliza para trás ao longo da fita molde de DNA. O recuo solta o grupo 3′-OH da molécula de RNA do sítio ativo da RNA polimerase e paralisa temporariamente a síntese de RNA. As células usam vários mecanismos para minimizar o recuo, incluindo as proteínas que rompem o RNA recuado no sítio ativo, gerando um novo 3′-OH ao qual novos nucleotídios podem ser adicionados. Nas células bacterianas, a tradução do mRNA pelos ribossomos quase acompanha a transcrição (ver *Capítulo 15*), e os ribossomos movendo-se ao longo do mRNA no sentido 5′ → 3′ também evitam o recuo da RNA polimerase na extremidade 3′ do mRNA.

As pausas temporárias na transcrição são importantes na coordenação da transcrição e da tradução nas bactérias, assim como na coordenação do processamento do RNA nos eucariotos. A pausa também afeta as taxas de síntese de RNA. Às vezes, uma pausa pode ser estabilizada por sequências no DNA que levam ao término da transcrição (ver, a seguir, *Término*).

Exatidão da transcrição. Embora a RNA polimerase seja muito exata na incorporação dos nucleotídios na cadeia de RNA crescente, ocasionalmente ocorrem erros. A pesquisa mostrou que a RNA polimerase faz um tipo de revisão durante a transcrição. Quando essa enzima incorpora um nucleotídio que não pareia no molde de DNA, ela retorna e retira os dois últimos nucleotídios (incluindo aquele erroneamente incorporado) a partir da cadeia de RNA crescente. A RNA polimerase então segue adiante, transcrevendo o molde de DNA novamente.

Conceitos

A transcrição é iniciada no sítio de início, que, nas células bacterianas, é ajustado para ligar a RNA polimerase à sequência consenso do promotor. Não é necessário um *primer*. A transcrição ocorre dentro da bolha de transcrição. O DNA é desenrolado à frente da bolha e enrolado atrás dela. Ocorrem pausas frequentes no processo de transcrição.

Término

A RNA polimerase adiciona nucleotídios à extremidade 3′ da molécula de RNA crescente até ela transcrever um terminador. A maioria dos terminadores é encontrada *upstream* do sítio onde o término realmente ocorre. Dessa forma, a transcrição não para subitamente quando a polimerase alcança o terminador, como um carro para no sinal vermelho. Em vez disso, ela para após o terminador ser transcrito, como um carro para somente após passar sobre um quebra-molas. São necessários vários eventos de sobreposição no terminador para finalizar a transcrição: a RNA polimerase precisa parar de sintetizar RNA, a molécula de RNA tem de ser liberada da RNA polimerase, a molécula de RNA recém-sintetizada precisa se dissociar completamente do DNA, e a RNA polimerase precisa se soltar do molde de DNA.

As células bacterianas apresentam dois tipos principais de terminadores. Os **terminadores dependentes de Rho** podem levar ao término da transcrição apenas quando houver uma proteína auxiliar chamada de **fator Rho**. Os **terminadores independentes de Rho** (também conhecidos como terminadores intrínsecos) podem levar ao término da transcrição na ausência de Rho.

Terminadores dependentes de Rho. Os terminadores dependentes de Rho têm duas características. A primeira é o terminador propriamente dito, composto por sequências de DNA que fazem com que a RNA polimerase pause. A segunda é uma sequência de DNA que codifica um segmento de RNA *upstream* do terminador, normalmente rico em nucleotídios citosina e sem estruturas secundárias. Essa sequência é chamada de sítio de utilização de Rho (rut), e serve como sítio de ligação para a proteína Rho. Uma vez que Rho esteja ligado ao RNA, ele se move para a extremidade 3′, seguindo a RNA polimerase (**Figura 13.13**). Quando a RNA polimerase encontra o terminador, ela faz uma pausa, tornando possível que Rho seja alcançado. A proteína Rho tem atividade helicase, usada para abrir o híbrido RNA-DNA na bolha de transcrição, encerrando a transcrição.

Terminadores independentes de Rho. Os terminadores independentes de Rho, que compõem cerca de 50% de todos os terminadores nos procariotos, apresentam duas características em comum. Primeiro, eles contêm repetições invertidas –

sequências de nucleotídios em uma fita que estão invertidas e são complementares. Quando as repetições invertidas são transcritas para o RNA, forma-se uma estrutura secundária em grampo (**Figura 13.14**). Segundo, nos terminadores independentes de Rho, uma fita de 7 a 9 nucleotídios de adenina segue a segunda repetição invertida no DNA molde. Sua transcrição produz uma fita de nucleotídios de uracila após o grampo no RNA transcrito.

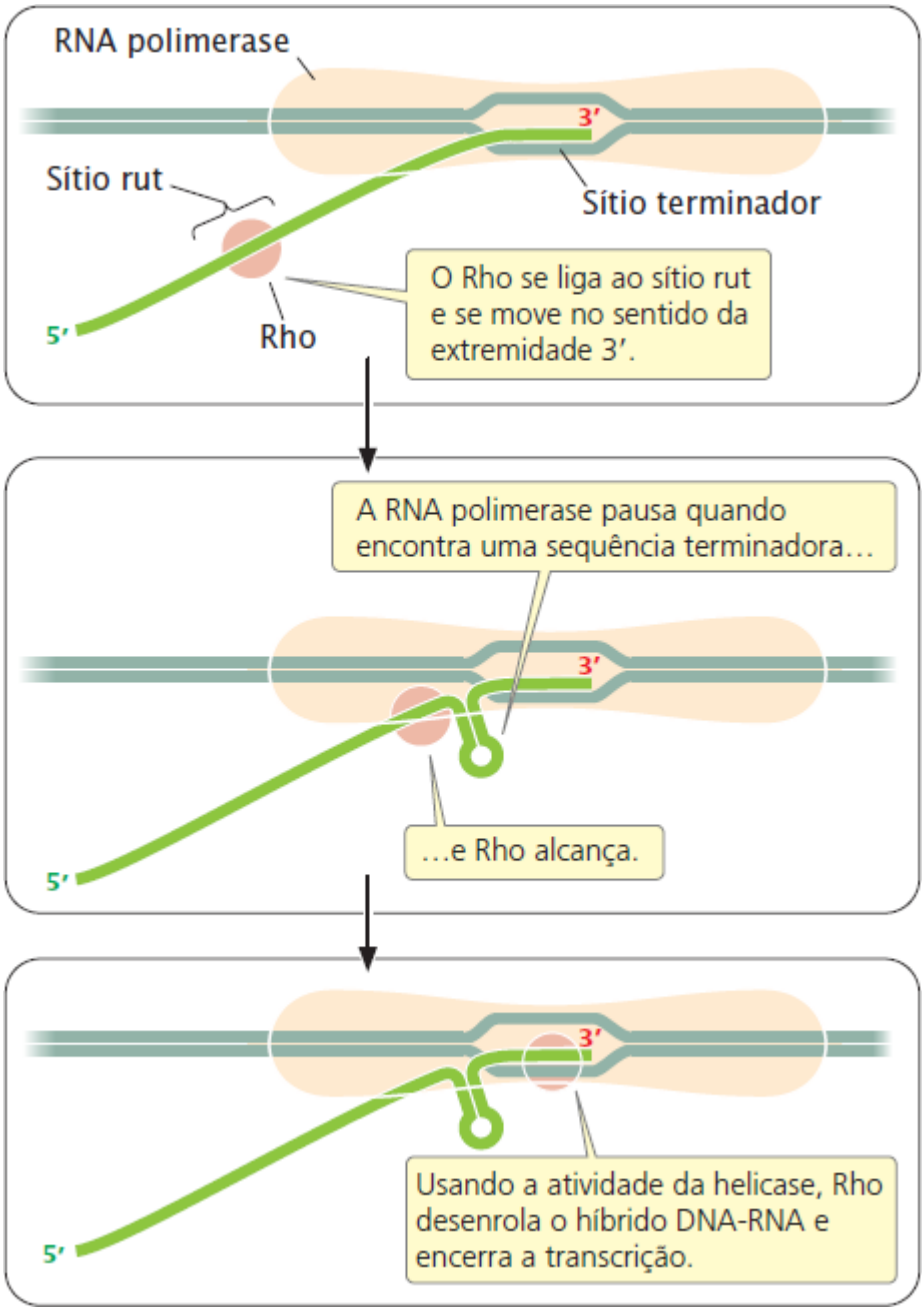


Figura 13.13 O término da transcrição em alguns genes bacterianos requer a proteína Rho.

A fita de uracilas na molécula de RNA faz com que a RNA polimerase pause, dando tempo para que a estrutura em grampo se forme. A evidência sugere que a formação do grampo desestabiliza o pareamento DNA-RNA, fazendo com que a molécula de RNA se separe de seu molde DNA. A separação pode ser facilitada pelos pareamentos das bases adenina-uracila, relativamente fracos se comparados com outros tipos de pareamentos de base. Quando o RNA transcrito se separa do molde, a síntese de RNA não pode mais continuar (ver Figura 13.13). Resolva o Problema 29

mRNA policistrônico. Nas bactérias, um grupo de genes é transcrito em uma molécula de RNA, que é chamada de **RNA policistrônico**. Assim, o RNA policistrônico é produzido quando um único terminador se encontra na extremidade de um grupo de vários genes que são transcritos juntos, em vez de cada gene ter seu próprio terminador. O mRNA policistrônico surge em alguns eucariotos, como *Caenorhabditis elegans*, mas é incomum.

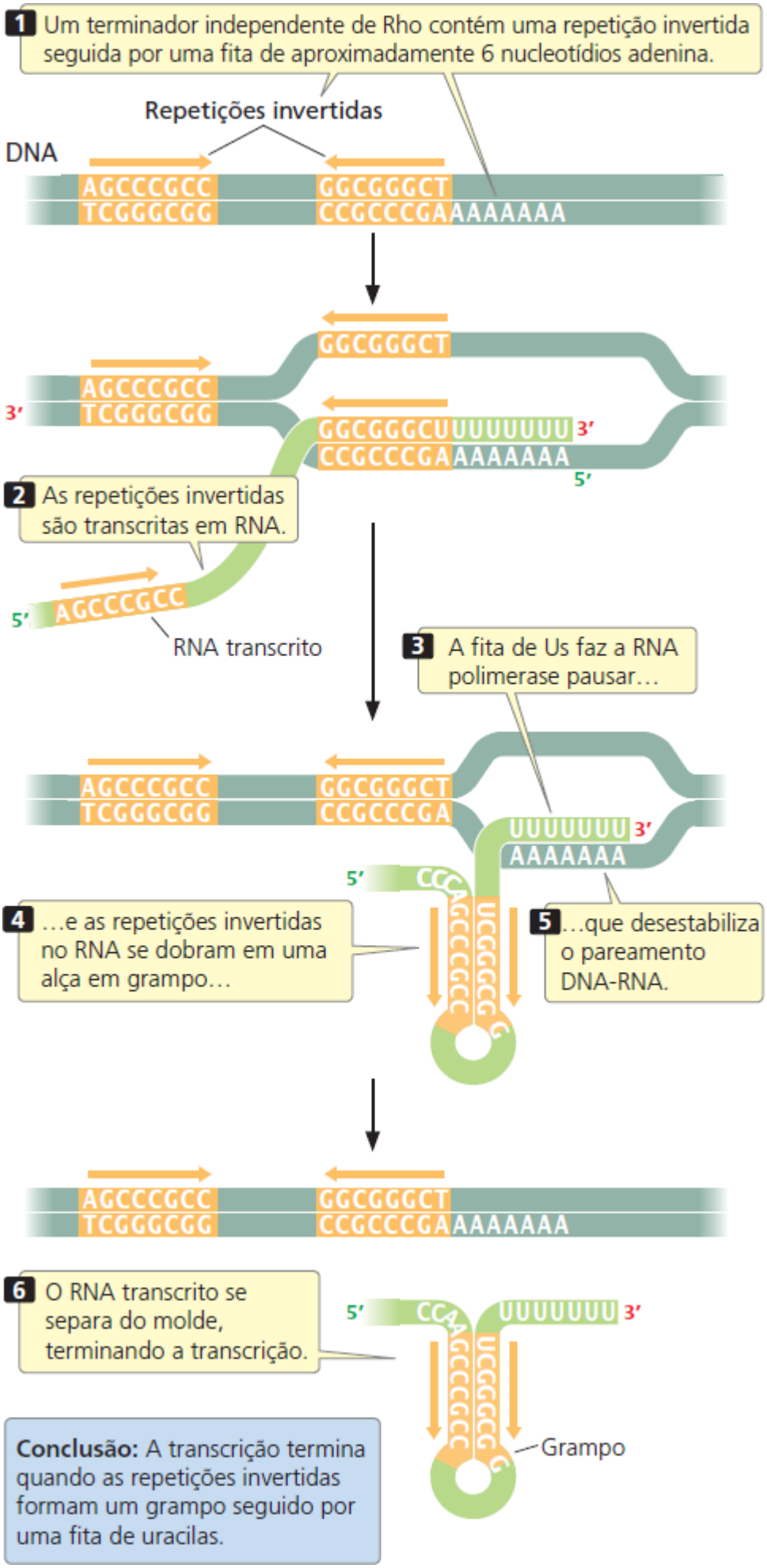


Figura 13.14 O término independente de Rho nas bactérias é um processo de várias etapas.

Conceitos

A transcrição termina após a RNA polimerase transcrever um terminador. As células bacterianas têm dois tipos de terminador: um terminador independente de Rho, o qual a RNA polimerase pode identificar por si só, e um terminador dependente de Rho, o qual o RNA polimerase pode identificar apenas com o auxílio da proteína Rho.

Checagem dos conceitos 6

Quais características são mais comumente encontradas nos terminadores independentes de Rho?

Conceitos conectantes

As regras básicas da transcrição

Antes de examinarmos o processo da transcrição eucariótica, resumiremos alguns dos princípios gerais da transcrição bacteriana.

1. A transcrição é um processo seletivo; apenas algumas partes do DNA são transcritas em qualquer momento.
2. O RNA é transcrito a partir do DNA de fita simples. Em um gene, apenas uma das duas fitas de DNA – a fita molde – é copiada em RNA.
3. Os trifosfatos de ribonucleosídios são usados como substratos na síntese do RNA. Dois grupos fosfato são rompidos a partir de um trifosfato de ribonucleosídio, e o nucleotídio resultante é unido ao grupo 3′-OH da fita de RNA crescente.
4. As moléculas de RNA são antiparalelas e complementares à fita molde de DNA. A transcrição é feita sempre no sentido 5′ → 3′, o significa que a molécula de RNA cresce para a extremidade 3′.
5. A transcrição depende da RNA polimerase – uma enzima complexa, multimérica. A RNA polimerase tem um cerne, que é capaz de sintetizar RNA, e outras subunidades que podem se unir temporariamente para executar outras funções.
6. Um fator sigma possibilita que o cerne da enzima da RNA polimerase se ligue a um promotor e inicie a transcrição.
7. Os promotores têm sequências curtas importantes na ligação da RNA polimerase ao DNA; essas sequências consenso estão intercaladas com os nucleotídios, que não têm participação conhecida na transcrição.
8. A RNA polimerase se liga ao DNA em um promotor, inicia a transcrição em um sítio de início do gene e termina a transcrição após um terminador ter sido transcrito.
9. As enzimas topoisomerases removem o superenrolamento que se desenvolve à frente e atrás da bolha de transcrição à medida que o DNA é desenrolado e enrolado durante a transcrição.

13.4 A transcrição eucariótica é semelhante à bacteriana, mas tem diferenças importantes

A transcrição nos eucariotos é semelhante à transcrição bacteriana, pois inclui início, alongamento e término, e os princípios básicos da transcrição já destacados também se aplicam à transcrição eucariótica. Entretanto, existem diferenças importantes. As células eucarióticas apresentam três diferentes RNA polimerases; cada uma transcreve uma classe diferente de RNA e identifica um tipo diferente de promotor. Assim, um promotor genérico não pode ser descrito para as células eucarióticas; de preferência, a descrição de um promotor depende se ele é identificado pela RNA polimerase I, II e III. Outra diferença é a natureza da identificação do promotor e do início. Muitas proteínas participam da ligação das RNA polimerases eucarióticas aos moldes de DNA, e os diferentes tipos de promotores exigem diferentes proteínas.

Transcrição e estrutura do nucleossomo

A transcrição requer que sequências no DNA sejam acessíveis à RNA polimerase e a outras proteínas. Entretanto, nas células eucarióticas, o DNA é complexado com as proteínas histonas na cromatina muito comprimida (ver Figura 11.3). Como as proteínas necessárias para a transcrição podem acessar o DNA eucariótico quando ele está complexado com as histonas?

A resposta é que a estrutura da cromatina é modificada antes da transcrição, de modo que o DNA esteja em uma configuração mais aberta e mais acessível para o maquinário de transcrição. Vários tipos de proteínas têm papéis na modificação da cromatina. As acetiltransferases adicionam grupos acetila a aminoácidos nas extremidades das proteínas histonas, o que desestabiliza a estrutura do nucleossomo e torna o DNA mais acessível. Outros tipos de modificação de histona também podem

afetar o empacotamento da cromatina. Além disso, as proteínas chamadas proteínas remodeladoras de cromatina podem se ligar à cromatina e deslocar os nucleossomos dos promotores e outras regiões importantes para a transcrição. Observaremos com mais detalhes o papel das mudanças na estrutura da cromatina associada à expressão gênica no Capítulo 17.

Conceitos

A iniciação da transcrição requer a modificação da estrutura da cromatina de modo que o DNA esteja acessível ao maquinário de transcrição.

Promotores

Uma diferença importante entre a transcrição nas bactérias e nos eucariotos é a existência de três RNA polimerases eucarióticas distintas, que identificam diferentes tipos de promotores. Nas células bacterianas, a holoenzima (RNA polimerase mais o fator sigma) identifica e se liga diretamente a sequências no promotor. Nas células eucarióticas, a identificação do promotor é feita pelas proteínas acessórias que se ligam ao promotor e então recrutam uma RNA polimerase específica (I, II ou III) para o promotor.

Uma classe de proteínas acessórias inclui os **fatores de transcrição gerais**, que junto com a RNA polimerase formam o **aparato basal de transcrição** – um grupo de proteínas que se reúnem próximo do sítio de início e são suficientes para iniciar os níveis mínimos de transcrição. Outra classe de proteínas acessórias é composta por **proteínas ativadoras de transcrição**, que se ligam a sequências específicas de DNA e elevam os níveis de transcrição ao estimular a montagem do aparato basal de transcrição no sítio de início.

Concentraremos nossa atenção nos promotores identificados pela RNA polimerase II, que transcreve os genes que codificam proteínas. Um promotor para um gene transcrito pela RNA polimerase II é composto por duas partes principais: um cerne do promotor e um promotor regulador.

Cerne do promotor. O **cerne do promotor** está localizado imediatamente *upstream* do gene (**Figura 13.15**) e é o sítio ao qual o aparato basal de transcrição se liga. O cerne do promotor inclui uma ou mais sequências consenso. Uma das sequências mais comuns é a **TATA box**, que tem a sequência consenso TATAAA e está localizada de −25 a −30 pb *upstream* do sítio de início. A Figura 13.15 mostra sequências consenso adicionais que podem ser encontradas nos cernes dos promotores dos genes transcritos pela RNA polimerase II. Essas sequências são reconhecidas pelos fatores de transcrição que se ligam a elas e servem como uma plataforma para a montagem do aparato basal de transcrição.

Promotor regulador. O **promotor regulador** está localizado imediatamente *upstream* do cerne do promotor. Podem ser encontradas várias sequências consenso diferentes nos promotores reguladores, e elas podem ser misturadas e unidas em diversas combinações. As proteínas ativadoras de transcrição se ligam a essas sequências e fazem contato direto ou indireto com o aparato basal de transcrição, e afetam a taxa na qual a transcrição é iniciada. As proteínas ativadoras de transcrição também regulam a transcrição ao se ligar a sequências mais distantes, chamadas de **acentuadores** (*enhancers*). O DNA entre o acentuador e o promotor faz uma alça, e as proteínas ativadoras de transcrição ligadas ao acentuador podem interagir com o maquinário basal de transcrição no cerne do promotor. Os acentuadores serão discutidos com mais detalhes no Capítulo 17.

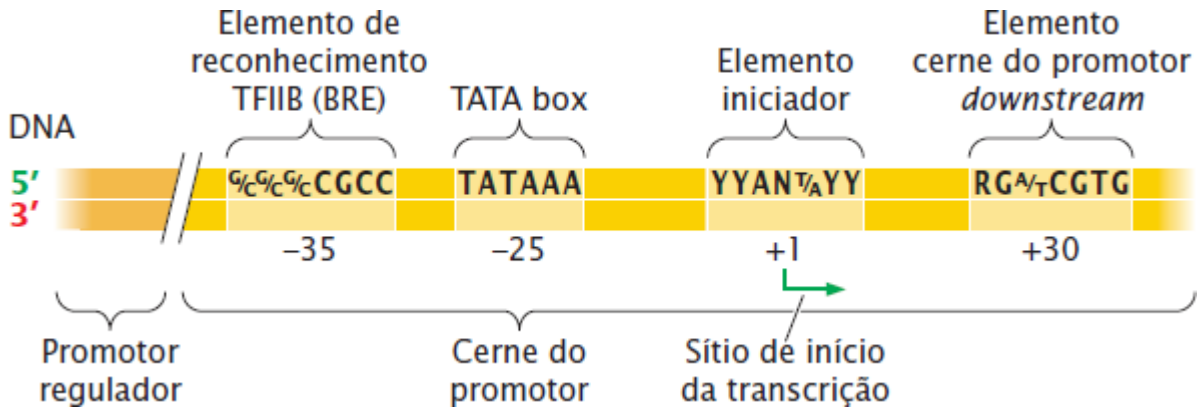


Figura 13.15 Os promotores dos genes transcritos pela RNA polimerase II consistem em um cerne do promotor e um promotor regulador que contêm as sequências consenso. Nem todas as sequências consenso apresentadas são encontradas em todos os promotores.

Promotores das polimerases I e III. As RNA polimerases I e III podem identificar promotores diferentes daqueles

identificados pela RNA polimerase II. Por exemplo, os promotores para os genes de rRNA e tRNA pequenos, transcritos pela RNA polimerase III, contêm **promotores internos** que estão *downstream* do sítio de início e são transcritos em RNA.

Conceitos

Os fatores de transcrição gerais e a RNA polimerase se montam no aparato basal de transcrição, que se liga ao DNA próximo ao sítio de início e é necessário para que a transcrição ocorra em níveis mínimos. Proteínas adicionais, chamadas de ativadoras de transcrição, ligam-se a outras sequências consenso nos promotores e nos acentuadores e afetam a taxa de transcrição.

Checagem dos conceitos 7

Qual é a diferença entre o cerne do promotor e o promotor regulador?

a. Apenas o cerne do promotor tem sequências consenso.
b. O promotor regulador está mais *upstream* do gene.
c. Os fatores de transcrição se ligam ao cerne do promotor; as proteínas ativadoras de transcrição se ligam ao promotor regulador.
d. Ambas as respostas b e c.

Iniciação

A transcrição nos eucariotos é iniciada pela montagem do maquinário de transcrição no promotor. Esse maquinário é composto por RNA polimerase II e uma série de fatores de transcrição que formam um complexo gigante com 50 ou mais polipeptídios. A montagem do maquinário de transcrição começa quando as proteínas reguladoras se ligam ao DNA próximo do promotor e modificam a estrutura da cromatina para que ocorra a transcrição. Essas e outras proteínas regulatórias, então, recrutam o aparato basal de transcrição para o cerne do promotor.

O aparato basal de transcrição é composto por RNA polimerase, uma série de fatores gerais de transcrição e um complexo de proteínas conhecido como mediador (**Figura 13.16**). Os fatores gerais de transcrição incluem TFIIA, TFIIB, TFIID, TFIIE, TFIIF e TFIIH, no qual TFII representa o fator de transcrição para RNA polimerase II, e a letra final indica o fator individual.

A RNA polimerase II e os fatores gerais de transcrição se reúnem no cerne do promotor, formando um complexo de pré-início análogo ao complexo observado no início bacteriano. Lembre-se de que, nas bactérias, o fator sigma identifica e se liga à sequência promotora. Nos eucariotos, a função do sigma é substituída pela função dos fatores gerais de transcrição. Uma primeira etapa é a ligação do TFIID a TATA box no molde de DNA. O TFIID é composto por pelo menos nove polipeptídios. Um deles é a **proteína ligadora de TATA** (TBP), que identifica e se liga à sequência consenso TATA. A proteína ligadora de TATA liga-se ao sulco menor e fixa-se ao DNA como um selim molecular (**Figura 13.17**), dobrando o DNA e desenrolando-o parcialmente. Outros fatores de transcrição se ligam a sequências consenso adicionais no cerne do promotor e a RNA polimerase, e posicionam a polimerase sobre o sítio de início da transcrição.

Conceitos

A transcrição é iniciada quando o aparato basal de transcrição, composto pela RNA polimerase e por fatores de transcrição, se monta no cerne do promotor e se torna um complexo aberto.

Checagem dos conceitos 8

Qual é o papel do TFIID no início da transcrição?

Após a RNA polimerase e os fatores de transcrição serem montados no cerne do promotor, ocorrem mudanças conformacionais tanto no DNA quanto na polimerase. Essas mudanças fazem com que de 11 a 15 pb de DNA ao redor do sítio de início da transcrição se separem, produzindo um DNA de fita simples que serve como molde para a transcrição.

O molde de DNA de fita simples é posicionado dentro do sítio ativo da RNA polimerase, criando uma estrutura chamada de complexo aberto. Após o complexo aberto ter sido criado, a síntese de RNA começa à medida que os grupos fosfato são separados dos trifosfatos de nucleosídio e os nucleotídios são unidos para formar uma molécula de RNA. Como na transcrição bacteriana, a RNA polimerase pode gerar e liberar várias moléculas de RNA pequenas na transcrição abortada antes que a polimerase inicie a síntese de uma molécula de RNA inteira. **Resolva o Problema 35**

Alongamento

Após cerca de 30 pb de RNA serem sintetizados, a RNA polimerase deixa o promotor e entra no estágio de alongamento da transcrição. Muitos dos fatores de transcrição são deixados para trás no promotor e podem servir para reiniciar rapidamente a transcrição com outra enzima RNA polimerase.

A estrutura molecular da RNA polimerase II eucariótica e como ela atua durante o alongamento foram revelados com o trabalho de Roger Kornberg e seus colegas, pelo que Kornberg recebeu o prêmio Nobel de química em 2006. A RNA polimerase mantém a bolha de transcrição durante o alongamento, no qual cerca de oito nucleotídios de RNA permanecem pareados com a fita molde de DNA. A dupla-hélice do DNA entra na fenda na polimerase e é segurada por extensões semelhantes a pinças da enzima (**Figura 13.18**). As duas fitas de DNA são desenroladas e os nucleotídios de RNA complementares são adicionados à extremidade 3′ crescente da molécula de RNA. À medida que se afunila pela polimerase, o híbrido DNA-RNA encontra uma parede de aminoácidos e se dobra em ângulo quase reto; essa dobra posiciona a extremidade do híbrido DNA-RNA no sítio ativo da polimerase, e novos nucleotídios são adicionados à extremidade 3′ da molécula de RNA crescente. O RNA recém-sintetizado é separado do DNA e segue por outro sulco antes de deixar a polimerase.

Término

As três RNA polimerases eucarióticas usam diferentes mecanismos para o término da transcrição. A RNA polimerase I requer o fator de término semelhante ao fator Rho usado no término de alguns genes bacterianos. Diferente de Rho, que se liga à molécula de RNA recém-transcrita, o fator de término para a RNA polimerase I liga-se a uma sequência de DNA *downstream* do sítio de término.

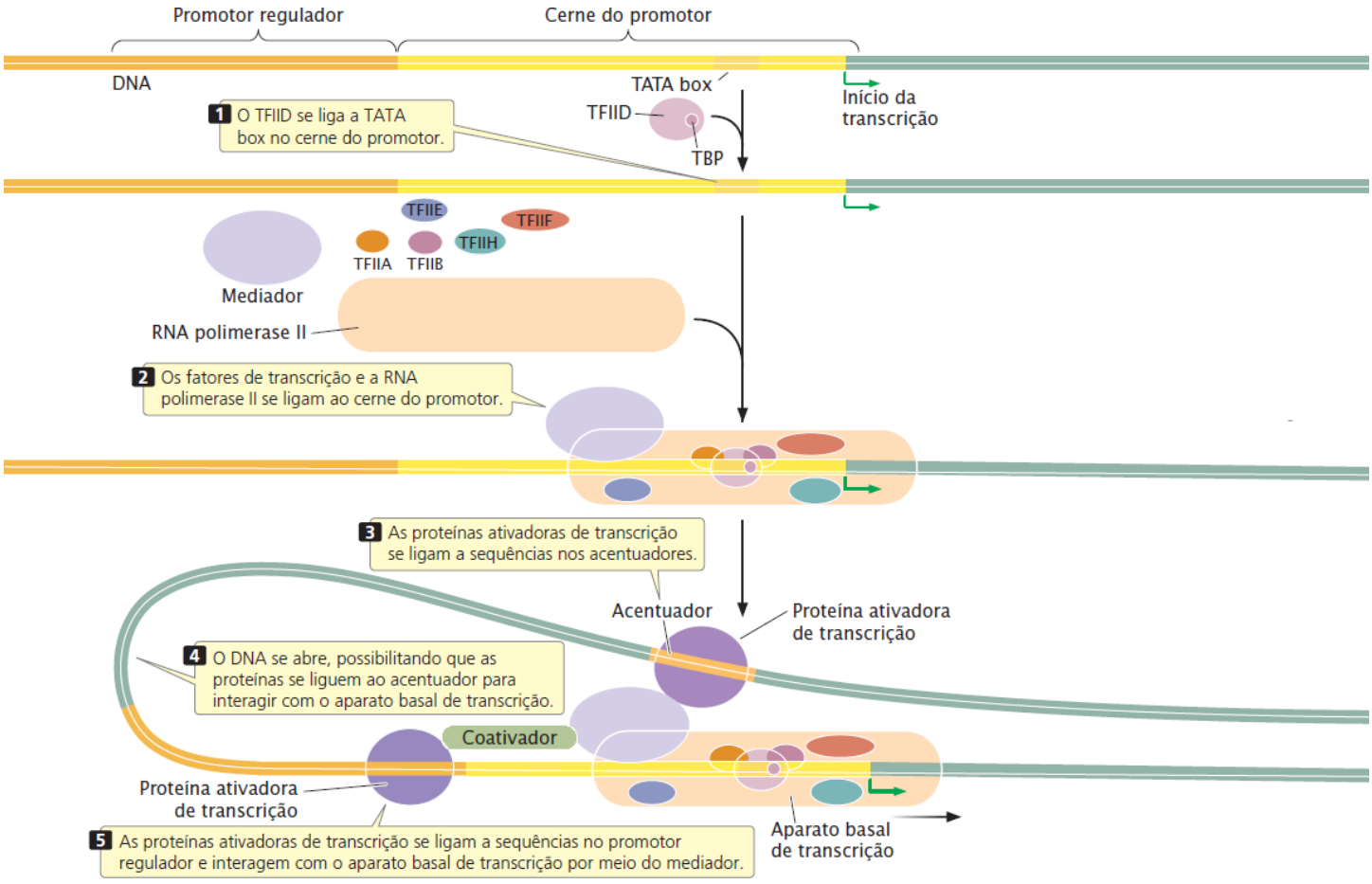


Figura 13.16 A transcrição é iniciada nos promotores da RNA polimerase II. A transcrição é iniciada quando o fator de transcrição TFIID se liga a TATA box, seguida pela ligação da holoenzima pré-montada contendo os fatores gerais de transcrição, RNA polimerase II e o mediador. TBP representa a proteína ligadora de TATA.

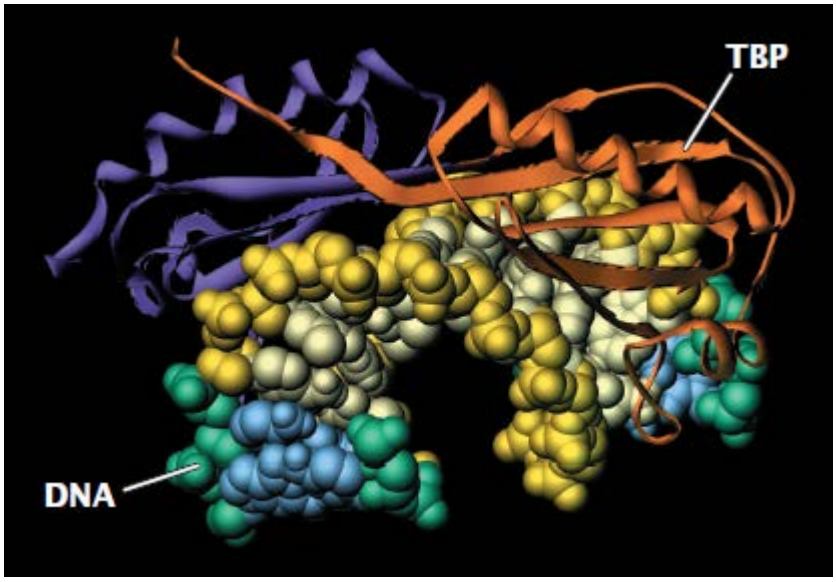


Figura 13.17 A proteína ligadora de TATA (TBP) liga-se no sulco menor do DNA, fixando-se à dupla-hélice de DNA como um selim.

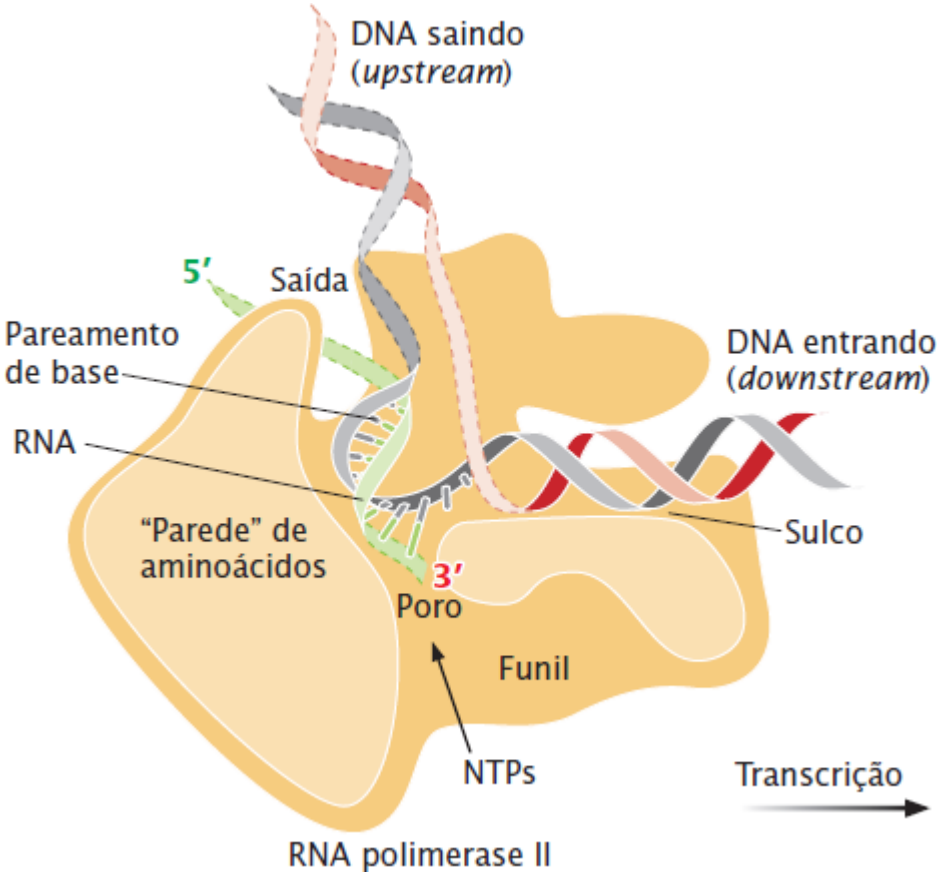


Figura 13.18 A estrutura da RNA polimerase II é uma fonte para compreender sua função. A dupla-hélice de DNA entra na polimerase por um sulco e se desenrola. O duplex DNA-RNA se dobra em um ângulo reto, que posiciona a extremidade 3′ do RNA no sítio ativo da enzima. Novos nucleotídios são adicionados à extremidade 3′ do RNA.

A RNA polimerase III termina a transcrição após transcrever uma sequência terminadora que produz uma fita de nucleotídios uracila na molécula de RNA, como a fita produzida pelos terminadores independentes de Rho das bactérias. Diferente dos terminadores independentes de Rho nas células bacterianas, entretanto, a RNA polimerase III não precisa que a estrutura em grampo anteceda a fita de uracilas.

O término da transcrição pela RNA polimerase II não ocorre em sequências específicas. Ao contrário, a RNA polimerase II continua a sintetizar RNA por centenas ou até milhares de nucleotídios depois da sequência de codificação necessária para produzir mRNA. Como observaremos no Capítulo 14, a extremidade do pré-mRNA é rompida em um sítio específico, designado por uma sequência consenso, enquanto a transcrição ainda ocorre na extremidade 3′ da molécula. A clivagem corta o pré-mRNA em dois fragmentos: o mRNA que vai codificar a proteína e outro segmento de RNA que tem sua extremidade 5′ se

arrastando para fora da RNA polimerase (**Figura 13.19**). Uma enzima (chamada Rat1 nas leveduras) se prende à extremidade 5′ desse RNA e se move para a extremidade 3′ onde a RNA polimerase continua a transcrição de RNA. A Rat1 é uma exonuclease 5′ → 3′ – uma enzima capaz de degradar o RNA no sentido 5′ → 3′. Como um torpedo guiado, a Rat1 se volta para a polimerase, "mastigando" o RNA à medida que se move. A transcrição termina quando a Rat1 alcança o maquinário de transcrição. Observe que esse mecanismo é semelhante ao do término dependente de Rho nas bactérias (ver Figura 13.13), exceto pelo fato de Rho não degradar a molécula de RNA.

Conceitos

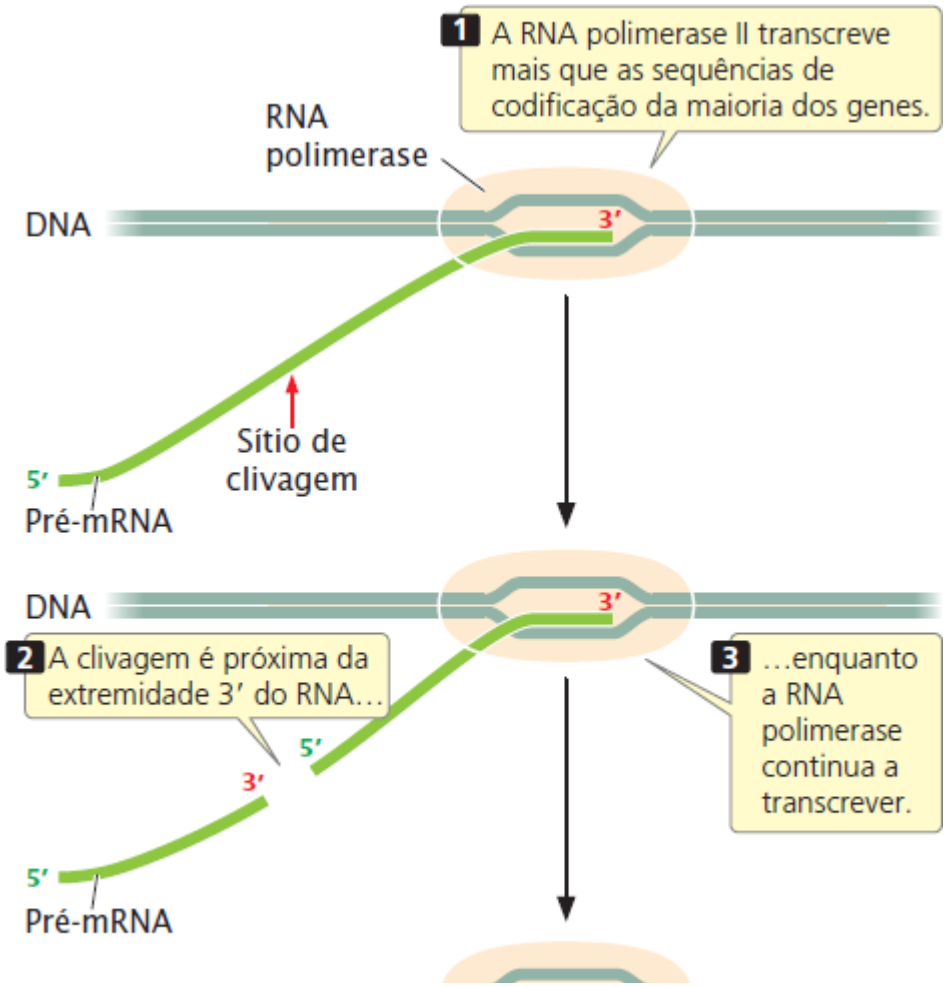
As três RNA polimerases eucarióticas diferentes usam mecanismos distintos de término. A transcrição em genes transcritos pela RNA polimerase II é encerrada quando uma enzima exonuclease se prende à extremidade 5′ rompida do RNA, desce o RNA e alcança a enzima polimerase.

Checagem dos conceitos 9

Quais são as semelhanças e as diferenças dos processos de término da RNA polimerase II e terminador dependente de Rho nas bactérias?

13.5 A transcrição em Archaea é mais semelhante à transcrição nos eucariotos do que à transcrição nas eubactérias

Há cerca de 2 a 3 bilhões de anos, a vida se dividia em três linhas de descendência evolutiva: as eubactérias, as archaea e os eucariotos (ver *Capítulo 2*). Embora as eubactérias e as archaea sejam superficialmente semelhantes – ambas são unicelulares e não apresentam núcleo –, os resultados dos estudos de suas sequências de DNA e de outras propriedades bioquímicas indicam que elas estão tão distantes uma da outra quanto estão dos eucariotos. A distinção evolutiva entre archaea, eubactérias e eucariotos é clara. No entanto, os eucariotos divergem do procarioto ancestral, com a última separação dos procariotos em eubactérias e archaea, ou as archaea e as eubactérias se dividiram primeiro, com os eucariotos evoluindo posteriormente a partir desses grupos?



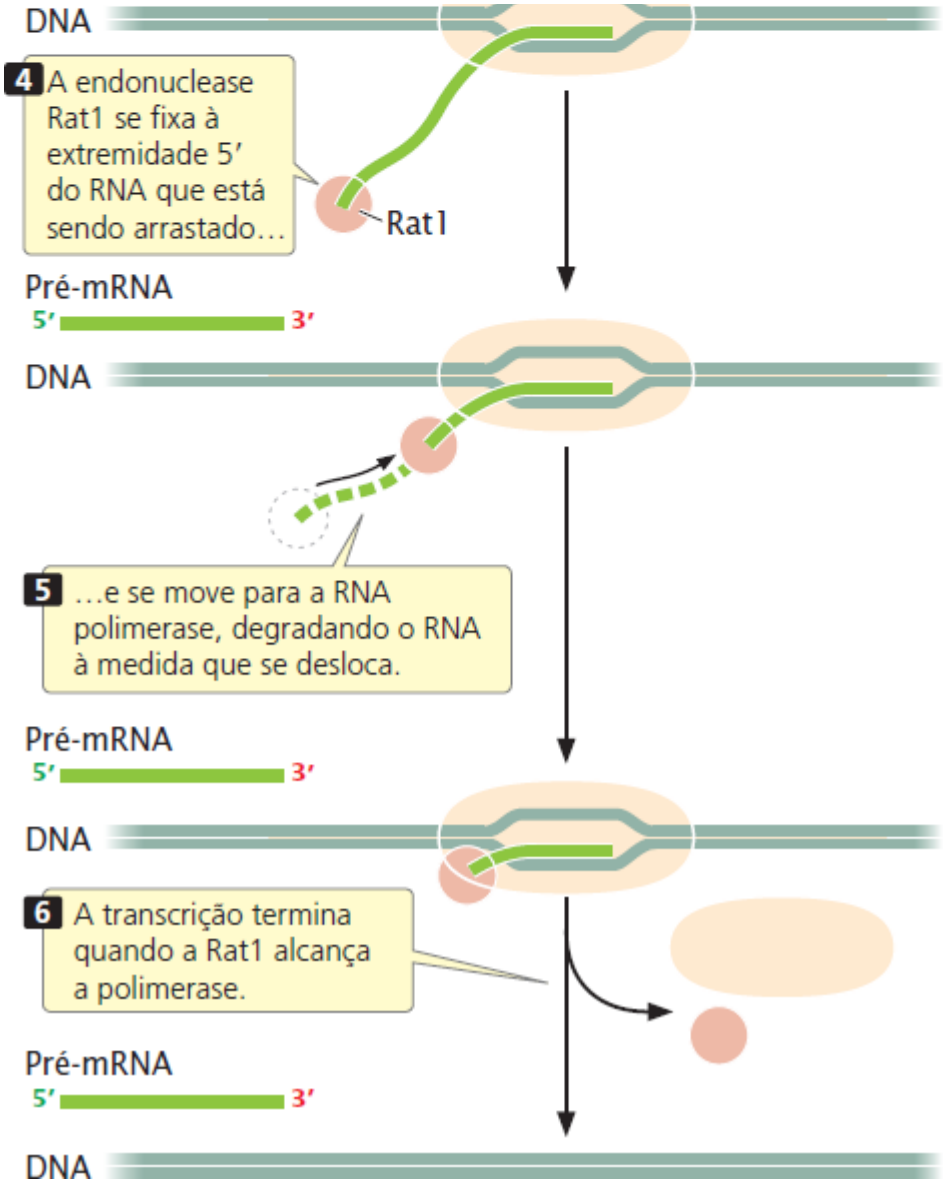


Figura 13.19 O término da transcrição pela RNA polimerase II requer a exonuclease Rat1. A clivagem do pré-mRNA produz uma extremidade 5′ à qual a Rat1 se fixa. A Rat1 degrada a molécula de RNA no sentido 5′ → 3′. A transcrição é paralisada quando a Rat1 alcança a polimerase.

Estudos da transcrição em eubactérias, archaea e eucariotos forneceram descobertas importantes sobre as relações evolutivas desses organismos. Archaea, como as eubactérias, têm uma única RNA polimerase, mas essa enzima é mais semelhante às RNA polimerases dos eucariotos. Como já foi discutido, a RNA polimerase bacteriana consiste em 5 subunidades, enquanto as RNA polimerases eucarióticas são muito mais complexas; por exemplo, a RNA polimerase II é composta por 12 subunidades. A RNA polimerase das archaea é da mesma forma complexa, com 11 ou mais subunidades. Além disso, a sequência de aminoácidos da RNA polimerase em archaea é semelhante à sequência de aminoácidos da RNA polimerase II eucariótica.

Os promotores de archaea apresentam uma sequência consenso semelhante a TATA box encontrada nos promotores eucarióticos. TATA box das archaea é encontrada aproximadamente 27 pb *upstream* do sítio de início da transcrição e, assim como nos eucariotos, ajuda a determinar a localização do sítio de início da transcrição. As archaea têm uma proteína ligadora de TATA (TBP), que é um fator de transcrição crítico encontrado em todas as três polimerases eucarióticas, mas não na RNA polimerase das eubactérias. A TBP se liga a TATA box em archaea com a ajuda de outro fator de transcrição, o TFIIB, que também é encontrado nos eucariotos, mas não nas eubactérias. Entretanto, alguns reguladores de transcrição encontrados em archaea são mais semelhantes aos encontrados nas bactérias, enfatizando que a transcrição em archaea não é totalmente eucariótica em natureza. Como os procariotos, as archaea não têm uma membrana nuclear, mas muitas espécies produzem histonas, que ajudam a compactar o DNA e a formar estruturas relacionadas ao nucleossomo.

Assim, a transcrição, um dos processos de vida mais básicos, tem muitas semelhanças nos eucariotos e nas archaea, sugerindo que esses dois grupos estão mais próximos um do outro do que a eubactérias. Esta conclusão é confirmada por outros dados, incluindo os obtidos a partir de uma comparação das sequências de genes.

Conceitos

O processo de transcrição na archaea tem muitas semelhanças com a transcrição nos eucariotos.

Resumo dos conceitos

- A vida primitiva usou RNA tanto como o carreador de informação genética quanto como catalisador biológico
- O RNA é um polímero, composto por nucleotídios unidos por meio de ligações fosfodiéster. Cada nucleotídio de RNA é constituído por um açúcar ribose, um fosfato e uma base. O RNA tem a base uracila e geralmente é de fita simples, o que possibilita a formação de estruturas secundárias
- As células têm várias classes diferentes de RNA. O RNA ribossômico é um componente do ribossomo, o RNA mensageiro carrega instruções de codificação para as proteínas e o RNA transportador ajuda a incorporar os aminoácidos em uma cadeia de polipeptídios
- O molde para a síntese de RNA é o DNA de fita simples. Na transcrição, a síntese de RNA é complementar e antiparalela à fita molde de DNA. Uma unidade de transcrição inclui um promotor, uma região codificadora de RNA e um terminador
- Os substratos para a síntese de RNA são trifosfatos de ribonucleosídio
- A RNA polimerase nas células bacterianas consiste em um cerne da enzima, que catalisa a adição de nucleotídios para uma molécula de RNA, e em outras subunidades. O fator sigma controla a ligação do cerne da enzima ao promotor
- As células eucarióticas apresentam muitas RNA polimerases diferentes
- O processo de transcrição é composto por três estágios: início, alongamento e término
- A transcrição começa no sítio de início, que é determinado pelas sequências consenso. Um segmento curto de DNA é desenrolado próximo do sítio de início, o RNA é sintetizado a partir de um molde de DNA de fita simples, e o DNA é enrolado na extremidade tardia da bolha de transcrição. As RNA polimerases são capazes de fazer revisão
- A síntese do RNA termina após a sequência terminadora ser transcrita. As células bacterianas têm dois tipos de terminadores: os terminadores independentes de Rho e os terminadores dependentes de Rho
- O início da transcrição nos eucariotos requer a modificação da estrutura da cromatina. Diferentes tipos de RNA polimerases nos eucariotos reconhecem diferentes tipos de promotores
- Para os genes transcritos pela RNA polimerase II, os fatores gerais de transcrição se ligam ao cerne do promotor e são parte do aparato basal de transcrição. As proteínas ativadoras de transcrição se ligam a sequências no promotor regulador e nos acentuadores e interagem com o aparato basal de transcrição no cerne do promotor
- As três RNA polimerases encontradas em todas as células eucarióticas usam diferentes mecanismos de término
- A transcrição em archaea tem muitas semelhanças com a transcrição dos eucariotos.

Termos importantes

Acentuador (*enhancer*)
Aparato basal de transcrição
Cerne da enzima
Cerne do promotor
CRISPR RNA (crRNA)
Elemento *upstream*
Fator de transcrição geral
Fator Rho
Fator sigma (σ)
Fita molde
Fita não molde
Holoenzima
Início abortado
MicroRNA (miRNA)
mRNA policistrônico
Promotor
Promotor interno
Promotor regulador
Proteína ativadora de transcrição
Proteína ligadora de TATA (TBP)
Região codificadora de RNA

Ribonucleoproteína nuclear pequena (snRNP)
Ribozima
RNA de interação com Piwi (piRNA)
RNA de interferência pequeno (siRNA)
RNA mensageiro (mRNA)
RNA nuclear pequeno (snRNA)
RNA nucleolar pequeno (snoRNA)
RNA polimerase
RNA polimerase I
RNA polimerase II
RNA polimerase III
RNA polimerase IV
RNA polimerase V
RNA pré-mensageiro (pré-mRNA)
RNA ribossômico (rRNA)
RNA transportador (tRNA)
Sequência consenso
Sequência consenso −10 (Pribnow box)
Sequência consenso −35
TATA box
Terminador
Terminador dependente de Rho
Terminador independente de Rho
Trifosfato de ribonucleosídio (rNTP)
Unidade de transcrição

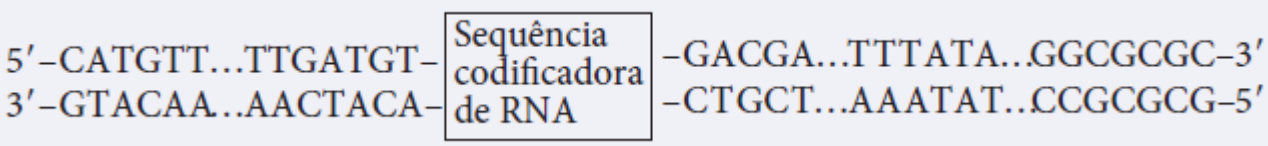
Respostas da Checagem dos conceitos

1. b.
2. A fita molde é a fita de DNA copiada em uma molécula de RNA, enquanto a fita não molde não é copiada.
3. d.
4. O fator sigma reconhece o promotor e controla a ligação da RNA polimerase ao promotor.
5. a.
6. As repetições invertidas seguidas por uma fita de nucleotídios de adenina.
7. d.
8. O TFIID se liga à TATA box e ajuda a centralizar a RNA polimerase sobre o sítio de início da transcrição.
9. Ambos os processos usam uma proteína que se liga à molécula de RNA e desce o RNA para a RNA polimerase. A diferença é que o Rho não degrada o RNA, enquanto a Rat1 degrada.

Problemas desenvolvidos

Problema 1

O digrama à direita representa uma sequência de nucleotídios que circundam uma sequência codificadora de RNA.



a. Esta sequência codificadora de RNA é de uma célula bacteriana ou de uma célula eucariótica? Como você pode saber?
b. Qual fita de DNA servirá como fita molde durante a transcrição da sequência codificadora do RNA?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

a. Se a sequência é oriunda de uma célula bacteriana ou eucariótica e por quê.
b. Qual fita é a fita molde.

Quais informações são fornecidas para solucionar o problema?

- As sequências de nucleotídios de ambas as fitas de DNA
- As extremidades 5′ e 3′ das fitas.

Para a solução deste problema, revise:
Molde, na Seção 13.2, Promotores bacterianos, na Seção 13.3, e Promotores, na Seção 13.4.

Etapas para a solução

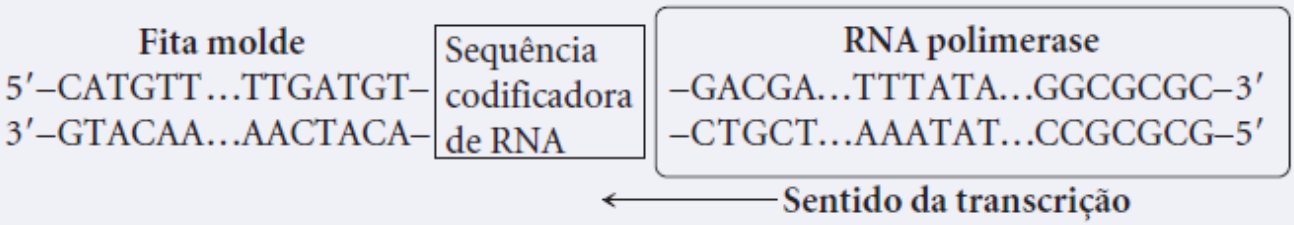
a. As células bacterianas e eucarióticas usam as mesmas bases de DNA (A, T, G e C), então as próprias bases não fornecem pistas sobre a origem da sequência. A sequência codificadora de RNA precisa ter um promotor, e as células bacterianas e eucarióticas têm sequências consenso diferentes encontradas em seus promotores; então, examinaremos as sequências quanto às sequências consenso familiares. Encontramos AAATAT na fita inferior à direita da sequência codificadora de RNA, que escrita de forma convencional (5′ à esquerda) é 5′–TATAAA–3′. Essa sequência é a TATA box encontrada na maioria dos promotores eucarióticos. Entretanto, essa sequência é bem semelhante à sequência consenso −10 (5′–TATAAT–3′), encontrada nos promotores bacterianos.

Dica: Revise as sequências consenso encontradas nos promotores bacterianos e eucarióticos nas Figuras 13.11 e 13.15 .

Mais longe à direita da fita inferior, também vemos 5′–GCGCGCC–3′, que é o elemento de reconhecimento de TFIIB (BRE, ver Figura 13.15) nos promotores de RNA polimerase II dos eucariotos. Não é encontrada sequência consenso semelhante nos promotores bacterianos; então, podemos ter a certeza de que essa sequência é um promotor eucariótico e uma sequência codificadora de RNA.

b. A TATA box e o BRE dos promotores da RNA polimerase II estão *upstream* das sequências codificadoras de RNA; assim, a RNA polimerase deve se ligar a essas sequências e então seguir *downstream*, transcrevendo a sequência codificadora de RNA. Desse modo, a RNA polimerase deve seguir da direita (*upstream*) para a esquerda (*downstream*). A molécula de RNA é sempre sintetizada no sentido 5′ → 3′ e é antiparalela à fita molde de DNA; então, a fita molde tem de ser lida no sentido 3′ → 5′. Se a enzima segue da direita para a esquerda e lê o molde no sentido 3′ → 5′, a fita superior tem de ser o molde, como apresentado no diagrama a seguir.

Lembrete: Durante a transcrição, a fita molde é lida no sentido 3′ → 5′.



Problema 2

Suponha que uma sequência consenso no promotor regulador de um gene eucariótico que codifica a enzima A foi detectada. Qual dos efeitos a seguir seria o resultado dessa deleção? Explique sua justificativa.

a. A enzima A teria uma sequência de aminoácidos diferente.
b. O mRNA para a enzima A seria curto e anormal.
c. A enzima A perderia alguns aminoácidos.
d. O mRNA para a enzima A seria transcrito, mas não traduzido.
e. A quantidade de mRNA transcrito seria afetada.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
A seleção do resultado (*a*, *b*, *c*, *d* ou *e*) que ocorreria quando fosse detectada uma sequência consenso no promotor regulador.

Quais informações são fornecidas para solucionar o problema?
A sequência consenso deletada está no promotor regulador.

Para a solução deste problema, revise:
Seção 13.4.

Etapas para a solução

A resposta correta é o item *e*. O promotor regulador tem os sítios de ligação para as proteínas ativadoras de transcrição. Essas sequências não são parte da sequência codificadora de RNA para a enzima A; então, a mutação não teria efeito no comprimento ou na sequência de aminoácidos da enzima, eliminando as respostas *a, b* e *c*. As proteínas ativadoras de transcrição se ligam ao promotor regulador e afetam a quantidade de transcrição que ocorre por meio das interações com o aparato basal de transcrição no cerne do promotor.

Questões de compreensão

Seção 13.1

1. Desenhe um nucleotídio de RNA e um nucleotídio de DNA, destacando as diferenças. Qual é a semelhança da estrutura do RNA em relação ao DNA? Qual a diferença?
2. Quais são as principais classes de RNA celular?
3. Por que o DNA é mais estável que o RNA?

Seção 13.2

4. Quais partes do DNA compõem uma unidade de transcrição? Desenhe uma típica unidade de transcrição bacteriana e identifique suas partes.
5. Qual é o substrato para a síntese do RNA? Como esse substrato é modificado e unido para produzir uma molécula de RNA?
6. Descreva a estrutura da holoenzima da RNA polimerase bacteriana.
7. Indique os nomes das RNA polimerases encontradas nas células eucarióticas e os tipos de RNA que elas transcrevem.

Seção 13.3

8. Quais são os três estágios básicos da transcrição? Descreva o que acontece em cada estágio.
9. Desenhe um típico promotor bacteriano e identifique as sequências consenso comuns.
10. Quais são os dois tipos básicos de terminadores encontrados nas células bacterianas? Descreva a estrutura de cada tipo.

Seção 13.4

11. Compare os papéis dos fatores gerais de transcrição e as proteínas ativadoras de transcrição.
12. Quais são as semelhanças e as diferenças entre a transcrição e a replicação?
13. Qual é a diferença entre a transcrição nas bactérias e nos eucariotos? Qual é a semelhança?

Questões e problemas aplicados

Seção 13.1

*14. Uma molécula de RNA tem as seguintes porcentagens de bases: A = 23%, U = 42%, C = 21% e G = 14%.
   a. Esta molécula de RNA tem fita simples ou fita dupla? Como você pode saber?
   b. Quais seriam as porcentagens de bases na fita molde do DNA que tem o gene para este RNA?

Seção 13.2

*15. O diagrama a seguir representa o DNA que é parte da sequência codificadora de uma unidade de transcrição. A fita inferior

é a fita molde. Indique a sequência encontrada na molécula de RNA transcrita a partir deste DNA e identifique as extremidades 5′ e 3′ do RNA.

5′-ATAGGCGATGCCA-3′
3′-TATCCGCTACGGT-5′ ← Fita molde

16. Para a molécula de RNA apresentada na **Figura 13.1 A**, escreva a sequência de bases nas fitas molde e não molde de DNA a partir das quais este RNA é transcrito. Marque as extremidades 5′ e 3′ de cada fita.
17. A sequência de nucleotídios a seguir foi encontrada em um molde de DNA de fita simples:

ATTGCCAGATCATCCCAATAGAT

Considere que a RNA polimerase segue ao longo deste molde da esquerda para direita.
   a. Qual extremidade do molde de DNA é 5′ e qual é 3′?
   b. Forneça a sequência e identifique as extremidades 5′ e 3′ do RNA copiado a partir deste molde.
18. As RNA polimerases fazem a transcrição em uma taxa muito menor que as DNA polimerases fazem a replicação. Por que a velocidade é mais importante na replicação do que na transcrição?
19. Considere que ocorre uma mutação no gene que codifica cada uma das RNA polimerases a seguir. Combine a mutação com os possíveis efeitos ao colocar a letra correta no espaço em branco. Existe mais de um efeito para cada polimerase mutante.

Mutação no gene que codifica	Efeitos
RNA polimerase I	_____
RNA polimerase II	_____
RNA polimerase III	_____

Possíveis efeitos

   a. tRNA não sintetizado.
   b. Algum RNA ribossômico não sintetizado.
   c. RNA ribossômico não processado.
   d. Pré-mRNA não processado.
   e. Algumas moléculas de mRNA não degradadas.
   f. Pré-mRNA não sintetizado.

Seção 13.3

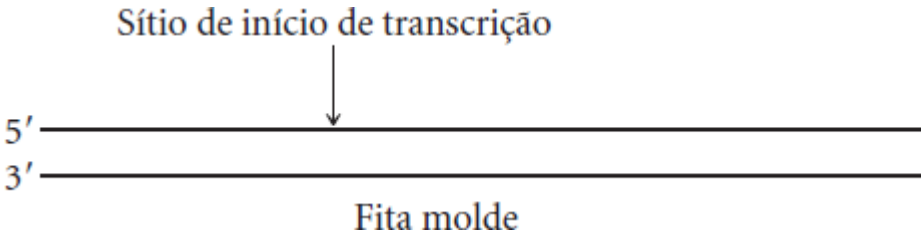
20. Forneça a sequência consenso para as *primeiras três* sequências reais apresentadas na **Figura 13.10**.
*21. Escreva a sequência consenso para o conjunto a seguir de sequências de nucleotídios.

AGGAGTT
AGCTATT
TGCAATA
ACGAAAA
TCCTAAT
TGCAATT

22. Liste pelo menos cinco propriedades que as DNA polimerases e as RNA polimerases têm em comum. Liste pelo menos três diferenças.
23. A maioria das moléculas de RNA tem três grupos fosfato na extremidade 5′, mas as moléculas de DNA não têm nenhuma. Explique essa diferença.
*24. Escreva uma sequência hipotética de bases que pode ser encontrada nos primeiros 20 nucleotídios de um promotor de um gene bacteriano. Inclua ambas as fitas de DNA e identifique as extremidades 5′ e 3′ de ambas as fitas. Certifique-se de incluir o sítio de início para transcrição e qualquer sequência consenso encontrada no promotor.
*25. Qual seria o efeito mais provável de uma mutação nos seguintes locais em um gene da *E. coli*?
   a. −8
   b. −35
   c. −20

d. Sítio de início

26. Uma cepa de bactérias tem uma mutação sensível à temperatura no gene que codifica o fator sigma. As bactérias mutantes produzem um fator sigma incapaz de se ligar à RNA polimerase em temperaturas elevadas. Qual efeito essa mutação terá no processo de transcrição quando as bactérias forem expostas a temperaturas elevadas?
27. Na Figura 13.5, indique a localização dos promotores e terminadores para os genes *a*, *b* e *c*.
28. O diagrama a seguir representa uma unidade de transcrição em uma molécula de DNA.



a. Considere que esta molécula de DNA é bacteriana. Desenhe a localização aproximada do promotor e do terminador para esta unidade de transcrição.
b. Considere que esta molécula de DNA é eucariótica. Desenhe a localização aproximada de um promotor de RNA polimerase II.

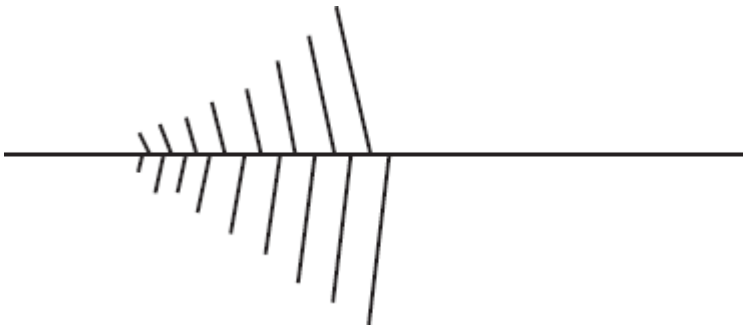
*29. Os nucleotídios de DNA a seguir foram encontrados próximo da extremidade de uma unidade de transcrição bacteriana.

3′-AGCATACAGCAGACCGTTGGTCTGAAAAAAGCATACA-5′

a. Marque o ponto no qual a transcrição terminará.
b. Este terminador é independente ou dependente de Rho?
c. Desenhe um diagrama do RNA que será transcrito a partir deste DNA, incluindo sua sequência de nucleotídios e qualquer estrutura secundária que surja.

30. Uma cepa de bactérias tem uma mutação sensível à temperatura no gene que codifica a subunidade Rho. O Rho não é funcional em altas temperaturas. Qual dos efeitos a seguir você espera ver se estas bactérias forem expostas a temperaturas elevadas? Explique sua justificativa para aceitar ou rejeitar cada uma das cinco opções.
a. A transcrição não ocorre.
b. Todas as moléculas de RNA são mais curtas que o normal.
c. Todas as moléculas de RNA são mais longas que o normal.
d. Algumas moléculas de RNA são mais longas que o normal.
e. O RNA é copiado de ambas as fitas de DNA.

31. O diagrama a seguir representa uma das estruturas semelhantes à árvore de Natal, como na **Figura 13.3**. Identifique as partes de *a* a *i* no diagrama.



a. Molécula de DNA.
b. Extremidades 5′ e 3′ da fita molde de DNA.
c. Pelo menos uma molécula de RNA.
d. As extremidades 5′ e 3′ de pelo menos uma molécula de RNA.
e. Sentido do movimento do aparato de transcrição na molécula de DNA.
f. Localização aproximada do promotor.
g. Possível localização de um terminador.
h. Sentidos *upstream* e *downstream*.
i. Moléculas de RNA polimerase (use pontos para representar essas moléculas).

32. Suponha que uma fita de nucleotídios A seguida pela repetição invertida em um terminador independente de Rho foi deletada, mas a repetição invertida ficou intacta. Como essa deleção afetará o término? O que acontecerá quando a RNA polimerase alcançar essa região?

Seção 13.4

33. O diagrama a seguir representa uma unidade de transcrição em uma molécula de DNA hipotética.

5′… TTGACA… TATAAT… 3′
3′… AACTGT… ATATTA… 5′

    a. Com base nas informações fornecidas, este DNA é bacteriano ou eucariótico?
    b. Se esta molécula de DNA for transcrita, qual fita será a fita molde e qual será a fita não molde?
    c. Onde, aproximadamente, será o sítio de início da transcrição?

34. Programadores, trabalhando em conjunto com geneticistas moleculares, desenvolveram programas de computador que podem identificar genes dentro de segmentos longos de sequências de DNA. Imagine que você está trabalhando com um programador nesse projeto. Com base no que você sabe sobre o processo de transcrição, quais sequências devem ser usadas para identificar o início e o final de um gene com o uso desse programa de computador?

*35. Por meio da engenharia genética, o geneticista faz uma mutação no gene que codifica TBP em células humanas cultivadas. Esta mutação destrói a capacidade do TBP em se ligar a TATA box. Preveja o efeito desta mutação nestas células.

36. Elaborados mecanismos de reparo estão associados à replicação para evitar mutações permanentes no DNA, sem que nenhum reparo semelhante esteja associado à transcrição. Você consegue pensar em um motivo para esta diferença na replicação e transcrição? (Dica: Pense nos efeitos relativos de uma mutação permanente em uma molécula de DNA comparada com uma mutação em uma molécula de RNA.)

Questões desafiadoras

Seção 13.3

37. Muitos genes nas bactérias e nos eucariotos têm várias sequências que potencialmente provocam pausas ou términos prematuros da transcrição. Contudo, a transcrição desses genes dentro de uma célula normalmente produz múltiplas moléculas de RNA com milhares de nucleotídios sem pausa ou término prematuro. Entretanto, quando um único ciclo de transcrição ocorre em tais moldes em um tubo teste, a síntese do RNA é frequentemente interrompida por pausas e términos prematuros, que reduzem a taxa na qual a transcrição ocorre e encurta o tamanho das moléculas de mRNA produzidas. A maioria das pausas e términos prematuros ocorre quando a RNA polimerase temporariamente retrocede (ou seja, dá marcha a ré) em um ou dois nucleotídios ao longo do DNA. Os achados experimentais demonstraram que a maioria dos atrasos de transcrição e términos prematuros desaparece se várias RNA polimerases forem simultaneamente transcritas da molécula de DNA. Proponha uma explicação para a transcrição mais rápida e o mRNA mais longo quando o DNA molde está sendo transcrito por múltiplas RNA polimerases.

Seção 13.4

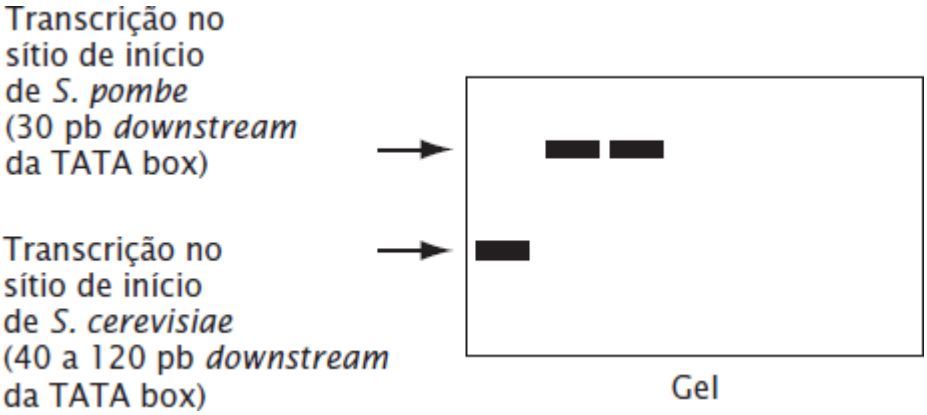
38. Os acentuadores são sequências que afetam o início da transcrição de genes que estão centenas ou milhares de nucleotídios distantes. As proteínas ativadoras de transcrição que se ligam aos acentuadores em geral interagem diretamente com os fatores de transcrição nos promotores, fazendo com que o DNA envolvido forme alças. Um acentuador do bacteriófago T4 não atua por meio de formação de alça no DNA (D. R. Herendeen *et al*. 1992. *Science* 256:1298-1303). Proponha alguns mecanismos adicionais (além das alças do DNA) pelos quais este acentuador possa afetar a transcrição em um gene a uma distância de milhares de nucleotídios.

*39. As localizações da TATA box em duas espécies de levedura, *Saccharomyces pombe* e *Saccharomyces cerevisiae* são muito diferentes. A TATA box da *S. pombe* está cerca de 30 nucleotídios *upstream* do sítio de início, semelhante à localização na maioria das células eucarióticas. Entretanto, a TATA box da *S. cerevisiae* está 40 a 120 nucleotídios *upstream* do sítio de início.

Para compreender melhor o que marca o sítio de início nestes organismos, os pesquisadores da Stanford University

conduziram uma série de experimentos para determinar quais componentes do aparato de transcrição destas duas espécies poderiam ser intercambiados (Y. Li *et al*. 1994. *Science* 263:805-807). Nesses experimentos, diferentes fatores de transcrição e RNA polimerase foram trocados em *S. pombe* e *S. cerevisiae* e foram observados os efeitos da troca no nível de síntese de RNA e no ponto de início da transcrição. Os resultados a partir de um conjunto de experimentos são apresentados no quadro a seguir. Componentes cTFIIB, cTFIIE, cTFIIF, cTFIIH são fatores de transcrição de *S. cerevisiae*. Componentes pTFIIB, pTFIIE, pTFIIF, pTFIIH são fatores de transcrição de *S. pombe*. Os componentes cPol II e pPol II são RNA polimerase II a partir de *S. cerevisiae* e *S. pombe*, respectivamente. O quadro indica se há componentes (+) ou se eles estão ausentes (−) no experimento. No gel anexo, a banda indica que o RNA foi produzido, e a posição da banda indica se ele está com o comprimento previsto quando a transposição inicia 30 pb *downstream* a partir da TATA box ou de 40 a 120 pb *downstream* a partir da TATA box.

	Experimento						
Componentes	**1**	**2**	**3**	**4**	**5**	**6**	**7**
cTFIIE	+	−	+	+	+	+	−
cTFIIH + cTFIIF	+	−	+	+	+	−	+
cTFIIB + cPol II	+	−	−	−	−	−	−
pPol II	−	+	+	+	−	+	+
pTFIIB	−	+	+	−	+	+	+
pTFIIE + pTFIIH + pTFIIF	−	+	−	−	−	−	−



a. Que conclusão podemos tirar sobre quais componentes determinam o sítio de início para transcrição a partir desses dados?
b. Que conclusões podemos tirar sobre as interações de diferentes componentes do aparato de transcrição?
c. Proponha um mecanismo para explicar por que o sítio de início para transcrição na *S. pombe* está cerca de 30 pb *downstream* da TATA box, enquanto o sítio de início para a transcrição na *S. cerevisiae* está 40 a 120 pb *downstream* da TATA box.

ANÁLISE DE DADOS

40. Glenn Croston e seus colegas estudaram a relação entre a estrutura da cromatina e a atividade da transcrição. Em um conjunto de experimentos, eles mediram o nível da transcrição *in vitro* de um gene da *Drosophila* pela RNA polimerase II com o uso de DNA e várias combinações de histonas (G. E. Croston *et al*. 1991. *Science* 251:643-649).

Primeiro, eles mediram o nível de transcrição para o DNA exposto sem proteínas histonas associadas. Então, mediram o nível de transcrição após os octâmeros do nucleossomo (sem H1) serem adicionados ao DNA. A adição dos octâmeros fez com que o nível de transcrição caísse em 50%. Quando tanto os octâmeros de nucleossomo quanto as proteínas foram são adicionados ao DNA, a transcrição foi reprimida, caindo para menos de 1% do obtido com o DNA exposto, como demonstrado no quadro a seguir.

GAL4-VP16 é uma proteína que se liga ao DNA de alguns genes eucarióticos. Quando a GAL4-VP16 é adicionada ao DNA, o nível de transcrição da RNA polimerase II é muito elevado.

Tratamento	**Quantidade relativa de transcrição**
DNA exposto	100

DNA + octâmeros	50
DNA + octâmeros + H1	< 1
DNA + GAL4-VP16	1.000
DNA + octâmeros + GAL4-VP16	1.000
DNA + octâmeros + H1 + GAL4-VP16	1.000

Mesmo com a proteína H1, a GAL4-VP16 estimula os níveis elevados de transcrição.

Proponha um mecanismo sobre como a proteína H1 reprime a transcrição e como a GAL4-VP16 supera essa repressão. Explique como seu mecanismo proposto produziria os resultados obtidos nestes experimentos.

14

Moléculas de RNA e Processamento do RNA

Uma doença na família real

No dia 12 de agosto de 1904, o Czar Nicholas Romanov II, da Rússia, escreveu em seu diário: “Um grande dia para não ser esquecido, quando a misericórdia de Deus nos visitou”. Naquele dia, veio ao mundo Alexei, primeiro filho de Nicholas e herdeiro do trono russo.

Ao nascer, Alexei era um bebê grande e bonito, com cachos dourados e olhos azuis, mas com 6 semanas de vida ele começou a apresentar hemorragia espontânea no umbigo. O sangramento persistiu por vários dias e causou grande alarme. À medida que crescia e começava a andar, Alexei frequentemente tropeçava e caía, como todas as crianças. Mesmo leves arranhões sangravam muito, e pequenas contusões levavam a sangramento interno substancial. Com esse quadro, ficou claro que o menino tinha hemofilia.

Caracterizada por coagulação lenta e sangramento excessivo, a doença é causada pela mutação em um dos vários genes que codificam as proteínas envolvidas no processo de coagulação sanguínea. Nos portadores de hemofilia, pequenas lesões podem levar a perda de sangue potencialmente fatal; o sangramento espontâneo nas articulações desgasta os ossos, com consequente deformidade.

Alexei sofria de hemofilia clássica, um distúrbio genético ligado ao cromossomo X. Ele herdou o gene da hemofilia de sua mãe, Alexandra, que era portadora da mutação. A hipótese é que o gene se originou com a Rainha Vitória da Inglaterra (ver Figura 6.8). Dez descendentes masculinos da Rainha Vitória tinham hemofilia. Seis descendentes do sexo feminino, incluindo sua neta Alexandra (mãe de Alexei), eram portadoras.

Durante a infância, Alexei sofreu vários episódios graves de sangramento. Os médicos da família real não conseguiam ajudá-lo durante essas crises, pois não havia tratamento para interromper a perda de sangue. A revolução russa explodiu nesse período. Os bolcheviques capturaram o Czar e sua família e os mantiveram presos na cidade de Yekaterinburg. Na noite de 16 de julho de 1918, um esquadrão de fuzilamento executou a família real e seus criados, incluindo Alexei e suas quatro irmãs. Por muitos anos, os corpos da família do Czar permaneceram desaparecidos, mas seus esqueletos foram recuperados de duas covas fora dos arredores de Yekaterinburg. A comparação do DNA mitocondrial e do nuclear dos ossos com os descendentes dos parentes comprovou que a ossada era de fato da família real. Embora essa análise tenha determinado a identidade dos restos mortais, a natureza molecular da hemofilia na família real ainda era desconhecida.

Em 2009, os geneticistas analisaram o DNA a partir dos ossos da família do Czar para determinar a natureza genética da hemofilia. Nas pessoas com a hemofilia ligada ao X, é comum ocorrer mutação em um dos dois genes no cromossomo X – o gene para o fator VIII ou o gene para o fator IX da coagulação sanguínea.

Nas células eucarióticas, os genes são frequentemente interrompidos por sequências não codificadoras. As partes de um gene que codificam os aminoácidos de uma proteína são chamadas éxons, enquanto as sequências não codificadoras são chamadas de íntrons. Inicialmente, éxons e íntrons são transcritos em pré-mRNA, mas em um processo chamado de *splicing* de RNA (ou recomposição do RNA), os íntrons são posteriormente cortados e os éxons são “colados”. O exame de sequências de DNA dos genes para os dois fatores de coagulação em Alexandra não revelou mutações nos nucleotídios que codificam os aminoácidos, mas ocorreu uma mutação em um dos íntrons de seu gene para o fator IX. Essa mutação alterou o *splicing* dos éxons, criando um novo sinal de parada na sequência codificadora do mRNA e produzindo um fator IX de coagulação truncado, defeituoso. Tanto os alelos normais quanto os mutantes foram detectados no DNA de Alexandra, como é esperado de um portador heterozigoto. O DNA de Alexei tinha apenas o alelo mutante, responsável por sua hemofilia.

Como foi ilustrado pela hemofilia de Alexei, o processamento do RNA é muito importante para a síntese adequada das proteínas. Nas células eucarióticas, as moléculas de RNA são, frequentemente, muito modificadas após a transcrição: para os genes que codificam proteínas, um nucleotídio especial chamado cap é adicionado à extremidade 5′, uma cauda de nucleotídios adenina é adicionada à extremidade 3′, e os íntrons são retirados do meio. Os rRNAs e os tRNAs também são

modificados após a transcrição nos eucariotos e nos procariotos. No Capítulo 13, nos dedicamos à transcrição, o processo de síntese de moléculas de RNA. Neste capítulo examinaremos a função e o processamento do RNA.

Começamos observando cuidadosamente a natureza do gene. Então, examinamos o RNA mensageiro (mRNA), sua estrutura e como ele é modificado nos eucariotos após a transcrição. Depois analisamos o RNA transportador (tRNA), a molécula adaptada que forma a interface entre os aminoácidos e o mRNA na síntese das proteínas. A seguir, examinamos o RNA ribossômico (rRNA), a estrutura e a organização dos seus genes e como os rRNAs são processados. Por fim, analisaremos um grupo recém-descoberto de vários RNAs pequenos com papéis importantes em diversas funções biológicas.

À medida que exploramos o mundo do RNA e seu papel na função dos genes, veremos evidências de duas características importantes desse ácido nucleico. Primeiro, o RNA é muito versátil, tanto em sua estrutura quanto na bioquímica. Ele pode assumir várias estruturas secundárias, que fornecem a base para sua diversidade funcional. Segundo, o processamento e a função do RNA incluem interações entre duas ou mais moléculas de RNA.

14.1 Muitos genes têm estruturas complexas

O que é um gene? Como foi destacado no Capítulo 3, a definição de gene muda à medida que exploramos diferentes aspectos da hereditariedade. Um gene foi definido no Capítulo 3 como um fator herdado que determina uma característica. Essa definição pode parecer vaga porque diz apenas o que um gene faz, em vez de explicar o que ele é. Contudo, ela era adequada naquele momento porque nosso objetivo era saber como os genes influenciam a herança dos traços. Não precisávamos leva em conta a natureza física do gene para aprender as regras da herança.

O conhecimento da estrutura química do DNA e do processo de transcrição agora nos possibilita ter uma noção mais precisa sobre o que é um gene. O Capítulo 10 descreveu como a informação genética é codificada na sequência de bases do DNA: um gene é composto por um conjunto de nucleotídios de DNA. Mas como muitos nucleotídios constituem um gene? E como a informação nesses nucleotídios está organizada? Em 1902, Archibald Garrod sugeriu, corretamente, que os genes codificam as proteínas. Elas são feitas a partir de aminoácidos; então, um gene tem os nucleotídios que especificam os aminoácidos de uma proteína. Portanto, por muitos anos, a definição de trabalho de um gene era um conjunto de nucleotídios que especifica a sequência de aminoácidos de uma proteína. À medida que os geneticistas aprenderam mais sobre a estrutura dos genes, entretanto, ficou claro que esse conceito era uma simplificação exagerada.

Organização dos genes

O trabalho inicial sobre a estrutura do gene foi realizado em grande parte examinando-se as mutações nas bactérias e nos vírus. Essa pesquisa levou Francis Crick, em 1958, a propor que os genes e as proteínas são **colineares** – que existe uma correspondência direta entre a sequência de nucleotídios do DNA e a sequência de aminoácidos de uma proteína (**Figura 14.1**). O conceito de colinearidade sugere que o número de nucleotídios em um gene deve ser proporcional ao de aminoácidos na proteína codificada por esse gene. No geral, esse conceito é verdadeiro para os genes encontrados nas células bacterianas e em muitos vírus, embora tais genes sejam um pouco mais longos do que seria esperado se a colinearidade fosse estritamente aplicada, porque os mRNAs codificados pelos genes têm sequências nas suas extremidades que não especificam os aminoácidos. No início, também se considerava que os genes e as proteínas eucarióticas eram colineares, mas então surgiram indícios de que a estrutura do gene eucariótico poderia ser muito diferente. As células eucarióticas têm muito mais DNA do que o necessário para codificar as proteínas (ver *Capítulo 11*). Além disso, muitas moléculas de RNA grandes observadas no núcleo não existiam no citoplasma, sugerindo que os RNAs nucleares sofrem algum tipo de mudança antes de serem exportados para o citoplasma.

Na década de 1970, contudo, a maioria dos geneticistas ficou surpresa com o anúncio de que nem todos os genes são contínuos. Os pesquisadores observaram que quatro sequências codificadoras em um gene de um vírus eucariótico eram interrompidas por nucleotídios que não especificam aminoácidos. Essa descoberta foi feita quando o DNA viral foi hibridizado com o mRNA transcrito desse mesmo DNA viral, e a estrutura hibridizada foi examinada usando um microscópio eletrônico (**Figura 14.2**). O DNA era muito maior que o mRNA porque regiões do DNA formavam alças a partir das moléculas hibridizadas. Essas regiões continham nucleotídios no DNA que não existiam nos nucleotídios codificadores no mRNA. Muitos outros exemplos de genes interrompidos foram descobertos; ficou evidente que a maioria dos genes eucarióticos tem trechos de nucleotídios codificadores e não codificadores.

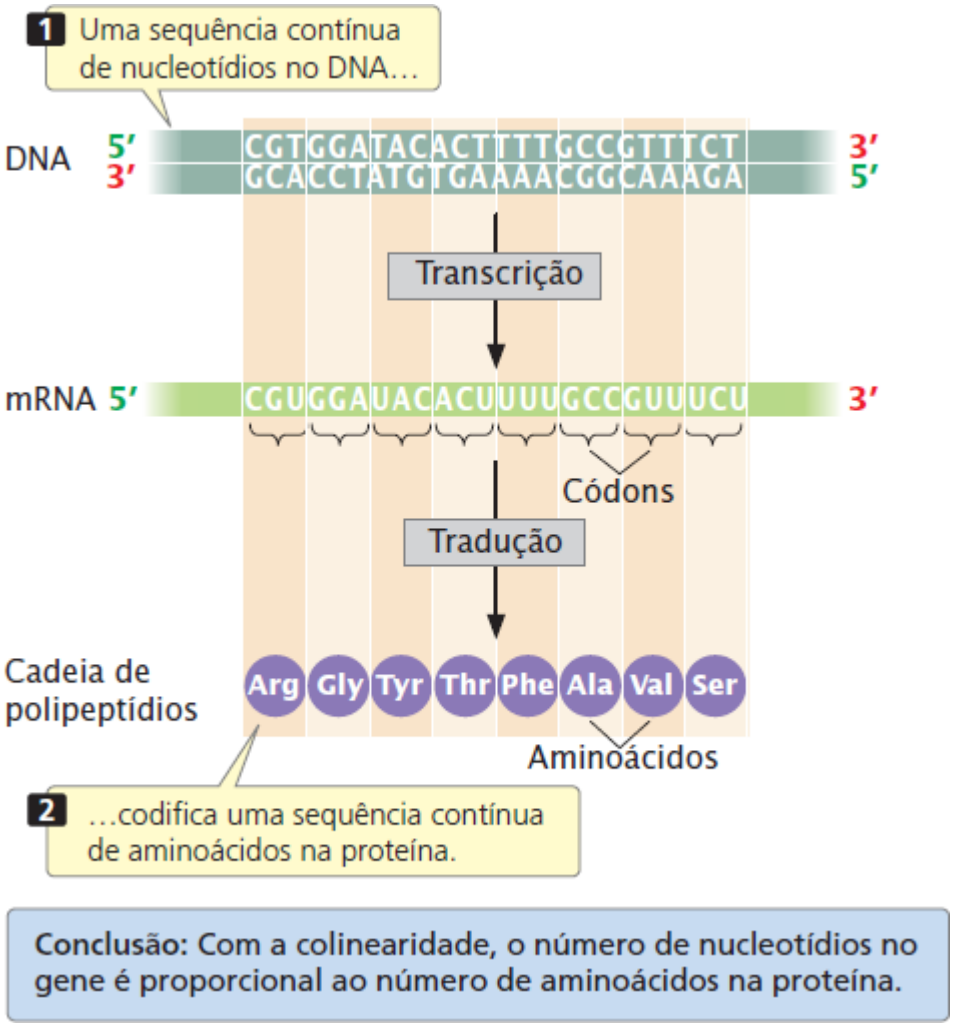
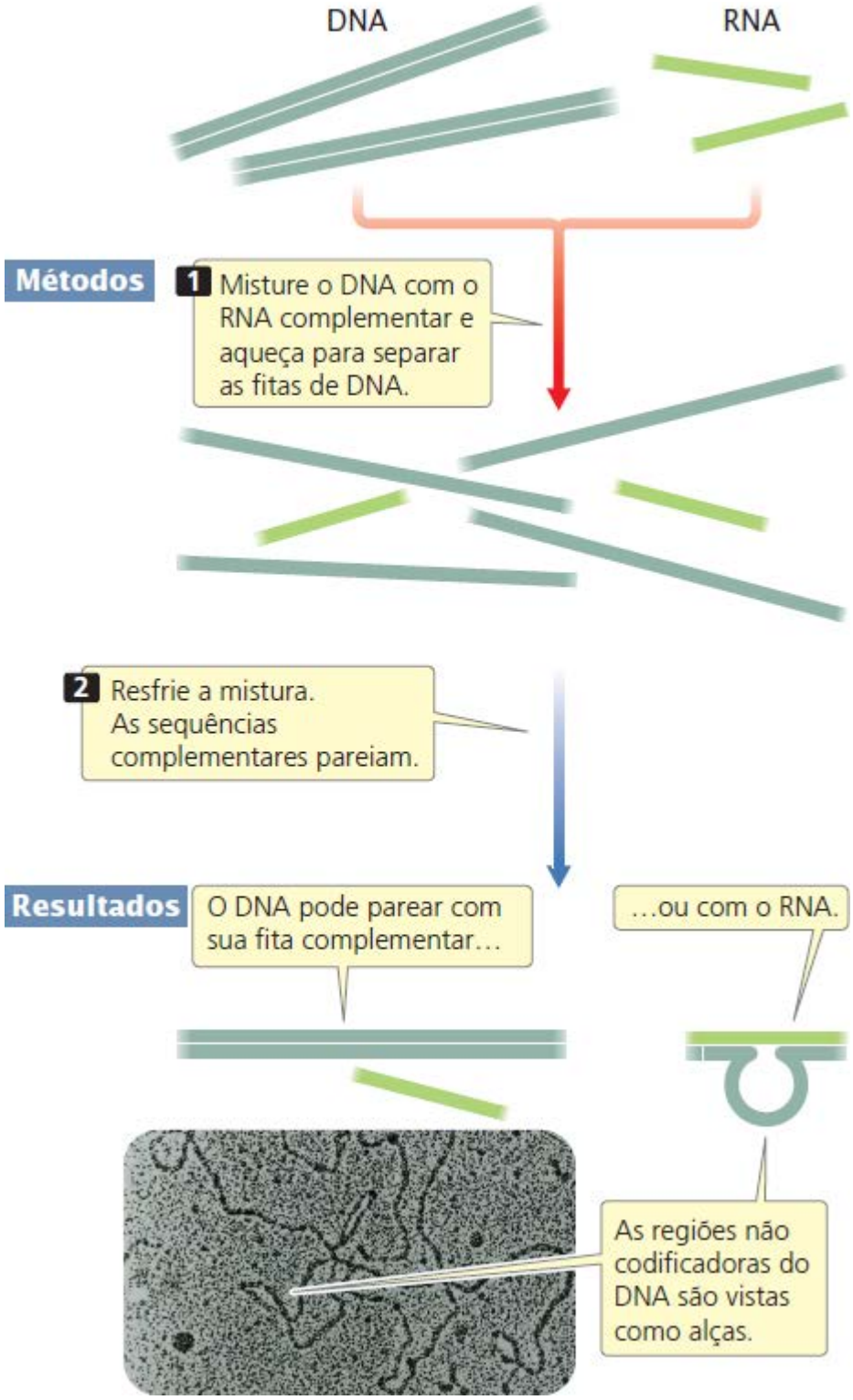


Figura 14.1 O conceito de colinearidade sugere que uma sequência contínua de nucleotídios no DNA codifica uma sequência contínua de aminoácidos em uma proteína.

Experimento

Pergunta: A sequência codificadora de um gene é sempre contínua?



Conclusão: As sequências codificadoras em um gene podem ser interrompidas por sequências não codificadoras.

Figura 14.2 A não colinearidade dos genes eucarióticos foi descoberta pela hibridização do DNA e do mRNA. (De Susan M. Berget, Claire Moore, and Phillip A. Sharp, Proc. Nati. Acad. Sci. USA 74(8), pp.3171-3175, Fig. 4g August 1977 Biochemistry.)

Conceitos

Quando uma sequência contínua de nucleotídios no DNA codifica uma sequência contínua de aminoácidos em uma proteína, os dois são chamados de colineares. Nos eucariotos, nem todos os genes são colineares com as proteínas que codificam.

Checagem dos conceitos 1

Que evidência indicou que os genes eucarióticos não são colineares com suas proteínas?

Íntrons

Muitos genes eucarióticos apresentam regiões codificadoras, chamadas **éxons,** e regiões não codificadoras, chamadas sequências intervenientes ou **íntrons**. Por exemplo, o gene que codifica a proteína ovalbumina tem oito éxons e sete íntrons; o gene para o citocromo *b* tem cinco éxons e quatro íntrons (**Figura 14.3**). O gene humano médio tem de oito a nove íntrons. Todos os íntrons e éxons são inicialmente transcritos em RNA, mas, após a transcrição, os íntrons são removidos por *splicing* e os éxons são unidos para gerar o RNA maduro.

Os íntrons são comuns nos genes eucarióticos e raros nos genes bacterianos. Por muitos anos após sua descoberta, acreditava-se que os íntrons estivessem totalmente ausentes dos genomas procarióticos, mas agora eles foram observados em archaea, nos bacteriófagos e até em algumas bactérias. Os íntrons são encontrados nos genes mitocondriais e cloroplásticos, assim como os genes nucleares dos eucariotos. Nos genomas eucarióticos, o tamanho e o número de íntrons parecem estar diretamente relacionados com a maior complexidade do organismo: genes de levedura têm apenas alguns íntrons curtos; os íntrons da *Drosophila* são mais longos e mais numerosos; e a maioria dos genes dos vertebrados é interrompida por longos íntrons. Todas as classes de genes eucarióticos – os que codificam rRNA, tRNA e proteínas – podem ter íntrons. O número e o tamanho de íntrons variam muito: alguns genes eucarióticos não apresentam íntrons, enquanto outros podem ter mais de 60; o comprimento do íntron varia de menos de 200 nucleotídios a mais de 50.000. Os íntrons tendem a ser mais longos que os éxons, e a maioria dos genes eucarióticos tem mais nucleotídios não codificadores do que codificadores. Por fim, a maioria dos íntrons não codifica proteínas: um íntron de um gene, em geral, não é um éxon para um gene diferente.

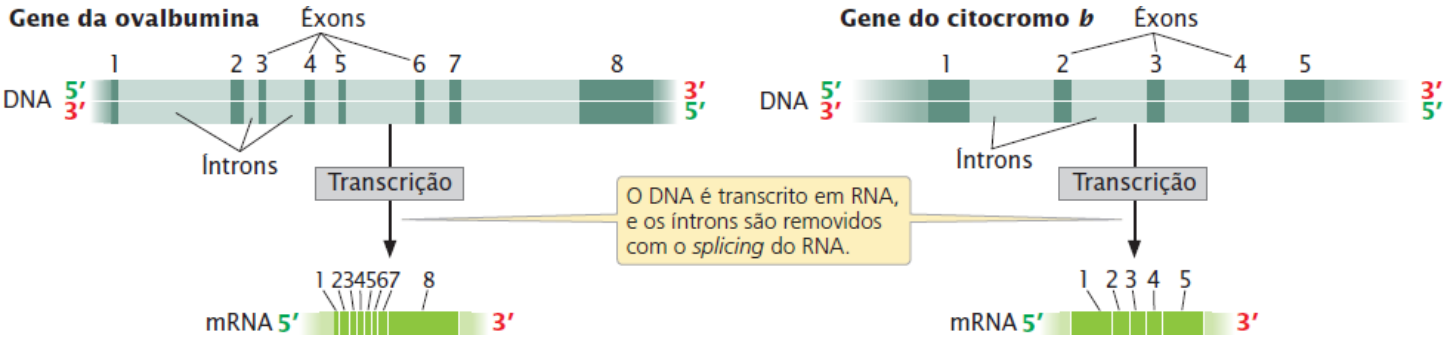


Figura 14.3 As sequências codificadoras de muitos genes eucarióticos são rompidas por íntrons não codificadores.

Há muito tempo os geneticistas discutem a origem evolutiva dos íntrons. Uma ideia, chamada de *hipótese tardia do íntron*, propõe que eles estavam ausentes nos organismos ancestrais, e foram adquiridos posteriormente pelos eucariotos. Outra ideia, chamada de *hipótese inicial do íntron*, sugere que os ancestrais primitivos das bactérias, de archaea e dos eucariotos tinham íntrons que, mais tarde, foram perdidos pelos procariotos e eucariotos simples. A evidência sugere que os íntrons foram perdidos e adquiridos com a evolução. Atualmente, muitos pesquisadores consideram que os primeiros eucariotos tinham íntrons, porque diferentes eucariotos apresentam íntrons nas mesmas posições em seus genes, sugerindo que esses íntrons podiam ser encontrados nos ancestrais de todos os eucariotos.

Existem quatro tipos principais de íntrons, diferenciados pela maneira como são removidos (**Quadro 14.1**). **Os íntrons do grupo I**, encontrados em alguns genes nas eubactérias, nos bacteriófagos e nos eucariotos, são *auto-splicing*: eles podem catalisar sua própria remoção. Os **íntrons do grupo II** são encontrados em alguns genes das mitocôndrias, dos cloroplastos, de archaea e de algumas eubactérias: eles também são *auto-splicing*, mas seu mecanismo de *splicing* é diferente dos íntrons do grupo I. Os **íntrons do pré-mRNA nuclear** são os mais bem estudados: eles incluem íntrons localizados nos genes que codificam as proteínas dos núcleos eucarióticos. O mecanismo de corte pelo qual esses íntrons são removidos é semelhante ao dos íntrons do grupo II, mas os íntrons nucleares não são *auto-splicing* – sua remoção requer snRNAs (discutido posteriormente neste capítulo) e várias proteínas. Os **íntrons de RNA transportador**, encontrados nos genes de tRNA de eubactérias, archaea e eucariotos, utilizam outro mecanismo de *splicing* que depende de enzimas para cortar e unir o RNA. Além desses principais grupos, existem vários outros tipos de íntrons.

Examinaremos com detalhes a química e o mecanismo de *splicing* do RNA mais adiante neste capítulo. No momento, vamos ter em mente as duas características gerais do processo de *splicing*: (1) o *splicing* de todos os íntrons de pré-mRNA ocorre no núcleo, e (2) a ordem dos éxons no DNA é mantida no RNA unido: as sequências codificadoras de um gene podem ser unidas, mas em geral não são misturadas. Resolva o Problema 19

Quadro 14.1 Principais tipos de íntrons.

Tipo de íntron	Localização	Mecanismo de *splicing*
Grupo I	Genes de eubactérias, bacteriófagos e eucariotos	*Auto-splicing*
Grupo II	Genes de eubactérias, archaea e organelas eucarióticas	*Auto-splicing*
Pré-mRNA nuclear	Genes que codificam proteínas no núcleo dos eucariotos	Spliceossômico
tRNA	Genes de tRNA de eubactérias, archaea e eucariotos	Enzimático

Nota: Existem vários tipos de íntrons menos importantes, incluindo o Grupo III, twintrons e íntrons de archaea.

Conceitos

Muitos genes eucarióticos têm éxons e íntrons. Ambos são transcritos em RNA, mas os íntrons são removidos posteriormente pelo processamento do RNA. O número e o tamanho dos íntrons variam entre os genes; eles são mais comuns em muitos genes eucarióticos, mas incomuns nos genes bacterianos.

Checagem dos conceitos 2

Quais são os quatro principais tipos de íntrons?

Conceito de gene revisitado

Como os íntrons afetam nosso conceito de gene? Definir um gene como uma sequência de nucleotídios que codifica aminoácidos em uma proteína já não parece ser adequado. Isso porque essa definição exclui tanto os íntrons que não especificam os aminoácidos quanto os nucleotídios que codificam as extremidades 5′ e 3′ de uma molécula de mRNA necessária para a tradução, mas que não codificam aminoácidos. Definir um gene nestes termos também exclui as sequências que codificam rRNA, tRNA e outros RNAs que não codificam proteínas. Considerando nossa atual compreensão da estrutura e da função do DNA, precisamos de uma definição mais precisa para o gene.

Muitos geneticistas ampliaram o conceito de gene para incluir todas as sequências no DNA que são transcritas em uma única molécula de RNA. Segundo essa definição, um gene inclui todos os éxons, os íntrons e as sequências no início e na extremidade do RNA que não são traduzidos em uma proteína. Ela também inclui as sequências de DNA que codificam rRNAs, tRNAs e outros tipos de RNA não mensageiro. Alguns geneticistas expandiram ainda mais a definição de gene para incluir a unidade de transcrição inteira, o promotor, a sequência codificadora de RNA e o terminador. Entretanto, nova evidência agora chama a atenção até para esta última definição. Pesquisa recente sugere que boa parte do genoma é transcrito em RNA, embora não esteja claro o que esse RNA faz. O que é certo é que o processo de transcrição é mais complexo do que se acreditava antes, e definir um gene como uma sequência que é transcrita em uma molécula de RNA não é tão claro como se pensava. Quanto mais aprendemos sobre a natureza da informação genética, mais evasiva se torna a definição de um gene.

Conceitos

A descoberta dos íntrons forçou uma reavaliação da definição do gene. Atualmente, um gene é definido como uma sequência de DNA que codifica uma molécula de RNA ou a sequência inteira de DNA necessária para transcrever e codificar uma molécula de RNA.

14.2 Os RNAs mensageiros que codificam as sequências de aminoácidos das proteínas são modificados após a transcrição nos eucariotos

Assim que o DNA foi identificado como a fonte da informação genética, ficou claro que ele não pode codificar diretamente as

proteínas. Nas células eucarióticas, o DNA se localiza no núcleo, mas a síntese da proteína ocorre no citoplasma. Os geneticistas identificaram que uma molécula adicional deve fazer parte na transferência de informação genética.

Os resultados dos estudos da infecção por bacteriófagos conduzidos no final dos anos 1950 e início dos anos 1960 destacaram o RNA como possível candidato a essa função de transferência. Os bacteriófagos injetam seu DNA nas células bacterianas, onde o DNA é replicado e grandes quantidades de proteína de fago são produzidas nos ribossomos bacterianos. No início de 1953, Alfred Hershey descobriu um tipo de RNA sintetizado rapidamente após a infecção pelo bacteriófago. Os achados de estudos posteriores evidenciaram que esse RNA de vida curta tinha uma composição de nucleotídios semelhante ao do DNA do fago, mas bem diferente do RNA bacteriano. Essas observações estavam coerentes com a ideia de que o RNA era copiado do DNA e de que esse RNA, então, direciona a síntese das proteínas.

Naquele momento, sabia-se que os ribossomos tinham participação, de *algum modo*, na síntese de proteínas, e boa parte do RNA na célula era conhecida na forma de ribossomos. Acreditava-se que os ribossomos fossem os agentes pelos quais a informação genética era deslocada para o citoplasma para a produção de proteínas. Usando a centrifugação por gradiente de densidade (ver Figura 12.2), Sydney Brenner, François Jacob e Matthew Meselson demonstraram, em 1961, que não era esse o caso. Eles mostraram que novos ribossomos *não* são produzidos durante a explosão de síntese de proteína que acompanha a infecção pelo fago (**Figura 14.4**). Os ribossomos não carreiam a informação genética necessária para produzir novas proteínas de fago.

Em outro experimento, François Gros e seus colegas infectaram células de *E. coli* com bacteriófagos enquanto adicionaram uracila marcada com radiação, que seria incorporada ao RNA recém-produzido do fago ao meio. Gros e seus colegas descobriram que o RNA recém-produzido de fato tinha vida curta, durava apenas alguns minutos, e estava associado aos ribossomos mas era diferente deles. Os cientistas concluíram que o RNA recém-sintetizado de vida curta carreia a informação genética para a estrutura de proteína ao ribossomo. Esse carreador recebeu o nome de *RNA mensageiro*.

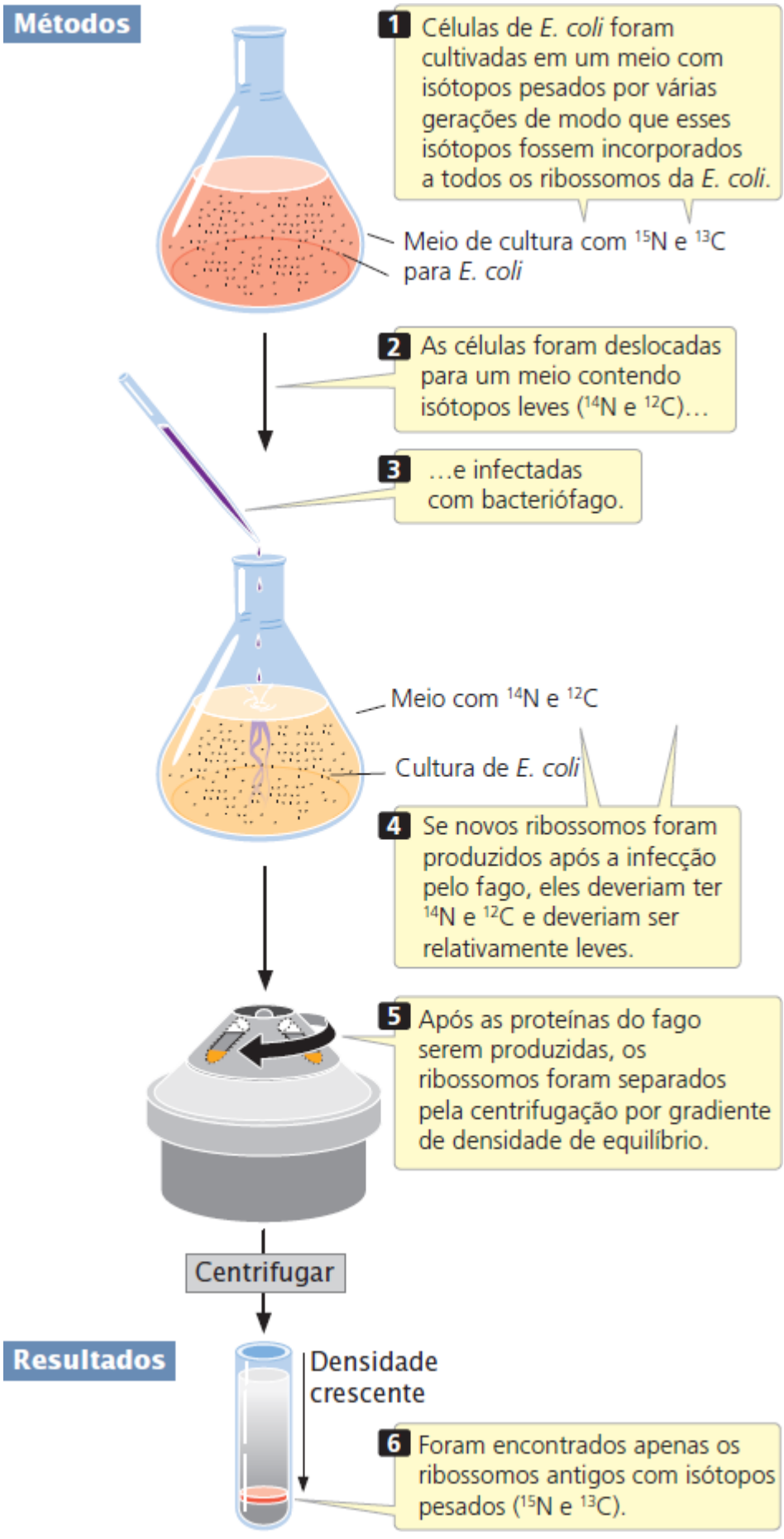
Estrutura do RNA mensageiro

O RNA mensageiro atua como molde para a síntese de proteína; ele carreia a informação genética do DNA para um ribossomo e ajuda a montar os aminoácidos na ordem correta. Nas bactérias, o mRNA é transcrito diretamente do DNA, mas nos eucariotos, um pré-mRNA (também chamado de transcrito primário) é transcrito primeiro, a partir do DNA, e então processado para gerar o mRNA maduro. Reservamos o termo mRNA para as moléculas de RNA que foram completamente processadas e estão prontas para serem traduzidas.

No mRNA, cada aminoácido em uma proteína é especificado por um conjunto de três nucleotídios chamado **códon**. Os mRNAs procarióticos e eucarióticos têm três regiões primárias (**Figura 14.5**). A **região 5′ não traduzida** (5′ UTR, às vezes chamada de líder) é uma sequência de nucleotídios na extremidade 5′ do mRNA que não codifica nenhum dos aminoácidos de uma proteína. No mRNA bacteriano, essa região tem uma sequência consenso (UAAGGAGGU) chamada de *sequência de* **Shine-Dalgarno**, que serve como sítio de ligação do cromossomo durante a tradução (ver *Capítulo 15*); são encontrados aproximadamente sete nucleotídios *upstream* do primeiro códon traduzido em um aminoácido (chamado de códon de início). Durante a tradução, a sequência de Shine-Dalgarno é complementar e pareia com as sequências encontradas em uma das moléculas de RNA que compõem o ribossomo. O mRNA eucariótico não tem sequência consenso equivalente na sua região 5′ não traduzida. Nas células eucarióticas, os ribossomos se ligam à extremidade 5′ modificada do mRNA, como será discutido posteriormente neste capítulo.

Experimento

Pergunta: Os ribossomos carreiam informações genéticas?



Conclusão: Os ribossomos não são produzidos na reprodução do fago e, portanto, não carreiam informação genética.

Figura 14.4 Brenner, Jacob e Meselson demonstraram que os ribossomos não carreiam informação genética.

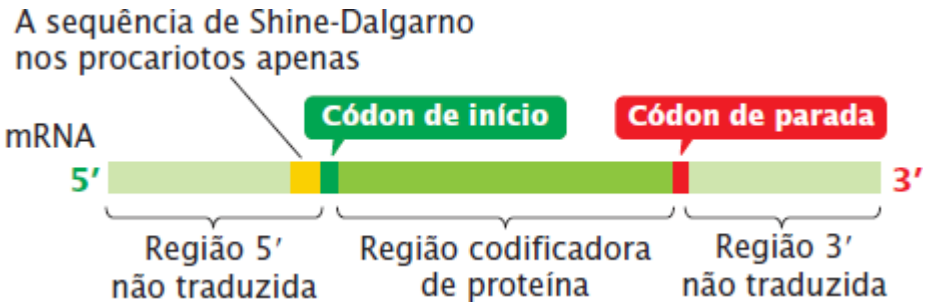


Figura 14.5 As três regiões primárias do mRNA maduro são a região 5′ não traduzida, a região codificadora de proteína e a região 3′ não traduzida.

A próxima seção do mRNA é a **região codificadora de proteína**, que inclui os códons que especificam a sequência de aminoácidos da proteína. A região codificadora de proteína começa com um códon de início e termina com um códon de parada. A última região do mRNA é a **região 3′ não traduzida** (3′ UTR, às vezes chamada de *trailer*), uma sequência de nucleotídios na extremidade 3′ do mRNA e não traduzida em proteína. A 3′ UTR afeta a estabilidade do mRNA e a tradução da sequência codificadora de proteína do mRNA.

Conceitos

As moléculas de RNA mensageiro têm três regiões principais: uma região 5′ não traduzida, uma região codificadora de proteína e uma região 3′ não traduzida. As regiões 5′ e 3′ não traduzidas não codificam aminoácidos de uma proteína, mas têm informações importantes para a tradução, a estabilidade do RNA e a regulação da expressão gênica.

Processamento do pré-mRNA

Nas células bacterianas, transcrição e tradução ocorrem simultaneamente; enquanto a extremidade 3′ de um mRNA está sendo transcrita, os ribossomos se prendem à sequência de Shine-Dalgarno próximo da extremidade 5′ e começam a tradução. Como a transcrição e a tradução estão acopladas, o mRNA bacteriano tem pouca oportunidade para ser modificado antes da síntese de proteínas. Nas células eucarióticas, ao contrário, a transcrição e a tradução são separadas no tempo e no espaço. A transcrição ocorre no núcleo, enquanto a tradução ocorre no citoplasma; tal separação fornece uma oportunidade para que o RNA eucariótico seja modificado antes de ser traduzido. Na verdade, o mRNA eucariótico é muito alterado após a transcrição. As mudanças ocorrem nas extremidades 3′ e 5′ e na seção codificadora de proteína da molécula de RNA (**Quadro 14.2**).

Quadro 14.2 Modificações pós-transcrição do pré-mRNA eucariótico.

Modificação	Função
Adição do cap 5′	Facilita a ligação do ribossomo à extremidade 5′ do mRNA, aumenta a estabilidade do mRNA e potencializa o *splicing* do RNA
Clivagem do 3′ e adição de uma cauda poli(A)	Aumenta a estabilidade do mRNA, facilita a ligação do ribossomo ao mRNA
Splicing do RNA	Remove os íntrons não codificadores do pré-mRNA, facilita a exportação do mRNA para o citoplasma, possibilita que múltiplas proteínas sejam produzidas por meio do *splicing* alternativo
Edição do RNA	Altera a sequência de nucleotídios do mRNA

Adição do cap 5′

Um tipo de modificação do pré-mRNA eucariótico é a adição de uma estrutura chamada **cap 5′**. O cap é um nucleotídio extra na extremidade 5′ do mRNA e grupos metila (CH_3) na base no nucleotídio recém-adicionado e no grupo 2′-OH do açúcar de um ou mais nucleotídios na extremidade 5′ (**Figura 14.6**). A adição do cap é feita rapidamente após o início da transcrição e, como será discutido com mais detalhes no Capítulo 15, atua no início da tradução. As proteínas ligadoras do cap o reconhecem e se ligam a ele; um ribossomo então se liga a essas proteínas e se move *downstream* ao longo do mRNA até que o códon de início

seja alcançado e a tradução comece. Apresentar um cap 5′ também aumenta a estabilidade do mRNA e influencia a remoção dos íntrons.

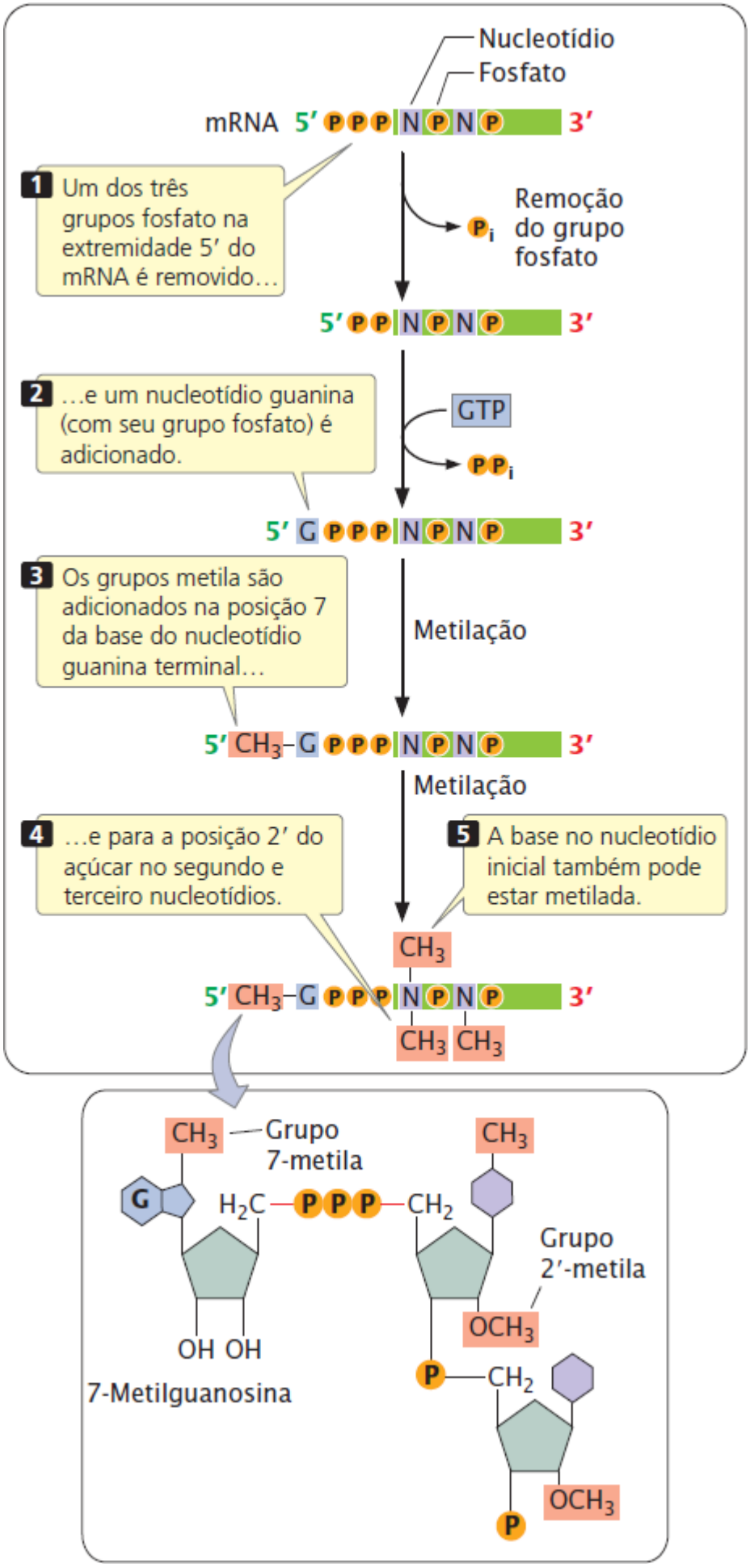


Figura 14.6 A maioria dos mRNAs eucarióticos apresenta cap 5′. O cap é um nucleotídio com 7-metilguanosina ligado ao pré-mRNA por uma ligação 5′-5′ única (apresentada em detalhes na imagem inferior).

Como foi observado na discussão sobre transcrição, no Capítulo 13, encontramos os três grupos fosfato na extremidade 5′ de todas as moléculas de RNA porque eles não são rompidos a partir do primeiro trifosfato de ribonucleosídio na reação de transcrição. A extremidade 5′ do pré-mRNA pode ser representada como 5′-pppNpNpN…, na qual "N" representa um ribonucleotídio e "p" representa um fosfato. Logo após o início da transcrição, um desses grupos fosfato é removido e um nucleotídio guanina é adicionado (ver Figura 14.6). Esse nucleotídio guanina é preso ao pré-mRNA por uma ligação 5′-5′ única, bem diferente da ligação fosfodiéster 5′-3′ usual que une todos os outros nucleotídios no RNA. Um ou mais grupos metila são então adicionados à extremidade 5′; o primeiro desses grupos metila é adicionado na posição 7 da base do nucleotídio guanina terminal, criando a base 7-metilguanina. A seguir, um grupo metila pode ser adicionado à posição 2′ do açúcar no segundo e terceiro nucleotídios (ver Figura 14.6). Em raros casos, grupos metila adicionais podem ser presos às bases do segundo e terceiro nucleotídios do pré-mRNA.

Várias enzimas diferentes participam da adição do cap 5′. A etapa inicial é feita por uma enzima associada à RNA polimerase II. Como nem a RNA polimerase I nem a RNA polimerase III apresentam essa enzima associada, as moléculas de RNA transcritas por essas polimerases (rRNAs, tRNAs e algumas snRNAs) não recebem o cap.

Adição da cauda poli(A)

Um segundo tipo de modificação para o mRNA eucariótico é a adição de 50 a 250 ou mais nucleotídios adenina na extremidade 3′, formando uma **cauda poli(A)**. Esses nucleotídios não são codificados no DNA, mas são adicionados após a transcrição (**Figura 14.7**) em um processo chamado de poliadenilação. A maioria dos genes eucarióticos transcrita pela RNA polimerase II é transcrita bem além da extremidade da sequência codificadora (ver *Capítulo 13*); a maior parte do material extra da extremidade 3′ é então rompida, e a cauda poli(A) é adicionada. Para algumas moléculas pré-mRNA, mais de 1.000 nucleotídios podem ser removidos da extremidade 3′ antes da poliadenilação.

O processamento da extremidade 3′ do pré-mRNA requer sequências, chamadas de sinal de poliadenilação, ambas *upstream* e *downstream* do sítio onde ocorre a clivagem. A sequência consenso AAUAAA está 11 a 30 nucleotídios *upstream* do sítio de clivagem (ver Figura 14.7) e determina o ponto onde ocorre a clivagem. Uma sequência rica em nucleotídios uracila (ou em nucleotídios guanina e uracila) está *downstream* do sítio de clivagem. Diversas proteínas participam da descoberta do sítio de clivagem e da remoção da extremidade 3′. Após o término da clivagem, os nucleotídios adenina são adicionados sem um molde à nova extremidade 3′, criando a cauda poli(A). Essa cauda confere estabilidade a vários mRNAs, aumentando o tempo durante o qual o mRNA permanece intacto e disponível para tradução antes de ser degradado por enzimas celulares. A estabilidade conferida pela cauda poli(A) depende das proteínas que se prendem à cauda e de seu comprimento. A cauda poli(A) também facilita a fixação do ribossomo ao mRNA e exerce papel na exportação do mRNA para o citoplasma.

Caudas poli(U) são adicionadas às extremidades 3′ de alguns mRNAs, microRNAs e pequenos RNAs nucleares. Embora a função das caudas poli(U) ainda esteja sendo investigada, a evidência sugere que em alguns mRNAs elas possam facilitar sua degradação. Resolva o Problema 29

Conceitos

Os pré-mRNAs eucarióticos são processados em suas extremidades 5′ e 3′. Um cap, composto por um nucleotídio modificado e vários grupos metila, é adicionado à extremidade 5′. O cap facilita a ligação a um ribossomo, aumenta a estabilidade do mRNA e pode afetar a remoção dos íntrons. O processamento da extremidade 3′ inclui a clivagem *downstream* de uma sequência consenso AAUAAA e a adição de uma cauda poli(A).

Checagem dos conceitos 3

Por que os mRNAs são cobertos, mas os tRNAs e rRNAs não são?

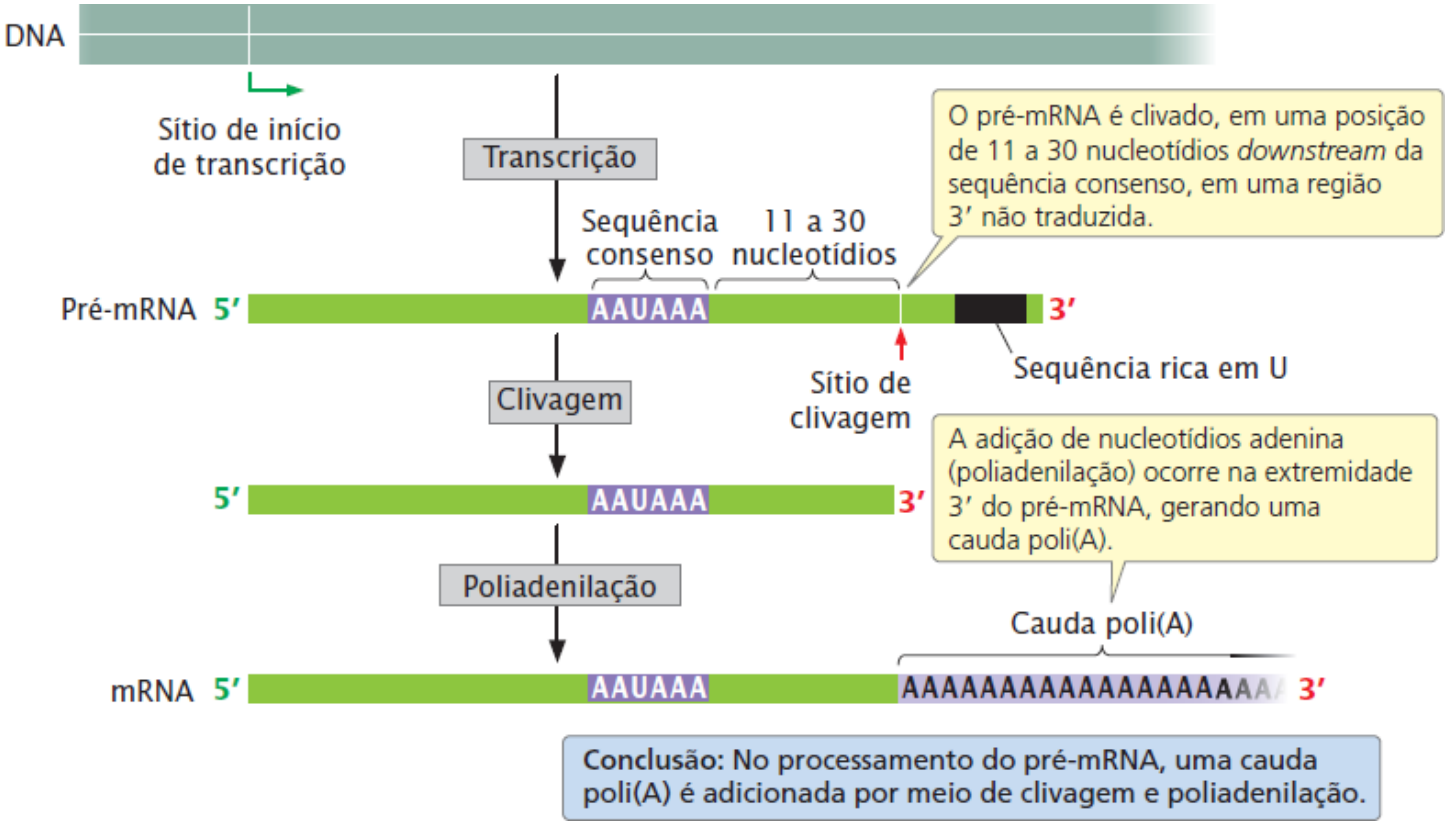


Figura 14.7 A maioria dos mRNAs eucarióticos tem cauda poli(A) 3′.

Splicing ou recomposição do RNA

Outro principal tipo de modificação do pré-mRNA eucariótico é a remoção de íntrons por ***splicing* ou recomposição do RNA**. Essa modificação ocorre no núcleo, antes que o RNA se desloque para o citoplasma.

Sequências consenso e spliceossomo. O *splicing* requer três sequências no íntron. Uma extremidade do íntron é chamada de **sítio de *splicing* 5′**, e a outra extremidade é o **sítio de *splicing* 3′** (**Figura 14.8**); esses sítios de união têm sequências consenso curtas. A maioria dos íntrons nos pré-mRNAs começa com GU e termina com AG, indicando que essas sequências têm um papel crucial no *splicing*. De fato, a mudança de um único nucleotídio em cada um desses sítios impede o *splicing*.

O terceiro nucleotídio importante para o *splicing* é o **ponto de ramificação**, que é um nucleotídio de adenina localizado 18 a 40 nucleotídios *upstream* do sítio de *splicing* 3′ (ver Figura 14.8). A sequência que circunda o ponto de ramificação não tem um consenso forte. A deleção ou mutação do nucleotídio adenina no ponto de ramificação evita o *splicing*.

O *splicing* ocorre dentro de uma grande estrutura chamada de **spliceossomo**, uma das maiores e mais complexa de todas as estruturas moleculares. O spliceossomo é composto por cinco moléculas de RNA e quase 300 proteínas. Os componentes do RNA são pequenos RNA nucleares (snRNAs, ver *Capítulo 13*), variando de 107 a 210 nucleotídios de comprimento; esses snRNAs se associam a proteínas para formar pequenas partículas de ribonucleoproteínas nucleares (snRNPs). Cada snRNP contém uma única molécula de snRNA e múltiplas proteínas. O spliceossomo é composto por cinco snRNPs (U1, U2, U4, U5 e U6) e algumas proteínas não associadas a um snRNA.

Conceitos

Os íntrons nos genes nucleares têm três sequências consenso críticas para o *splicing*: um sítio de *splicing* 5′, um sítio de *splicing* 3′ e um ponto de ramificação. O *splicing* do pré-mRNA ocorre dentro de um grande complexo chamado spliceossomo, composto por snRNAs e proteínas.

Checagem dos conceitos 4

Se um sítio de *splicing* sofresse uma mutação de modo que o *splicing* não ocorresse, qual seria o efeito no mRNA?

a. Ele seria menor que o normal.
b. Ele seria maior que o normal.

c. Ele teria o mesmo tamanho, mas codificaria uma proteína diferente.

Processo de splicing. Antes que o *splicing* ocorra, um íntron fica entre um éxon *upstream* (éxon 1) e um éxon *downstream* (éxon 2), como mostrado na **Figura 14.9**. O pré-mRNA é unido em duas etapas diferentes. Na primeira etapa do *splicing*, o pré-mRNA é cortado no sítio de *splicing* 5′. Esse corte libera o éxon 1 do íntrons, e a extremidade 5′ do íntron se prende ao ponto de ramificação; o íntron dobra-se sobre si mesmo, formando uma estrutura chamada de **lariat**. Nessa reação, o nucleotídio guanina na sequência consenso no sítio de *splicing* 5′ se liga com o nucleotídio adenina no ponto de ramificação por meio de uma reação de transesterificação. Nessa reação, a clivagem de 5′ e a formação do lariat ocorrem em uma única etapa. O resultado é que o grupo fosfato 5′ do nucleotídio guanina está agora preso ao grupo 2′-OH do nucleotídio adenina no ponto de ramificação (ver Figura 14.9).

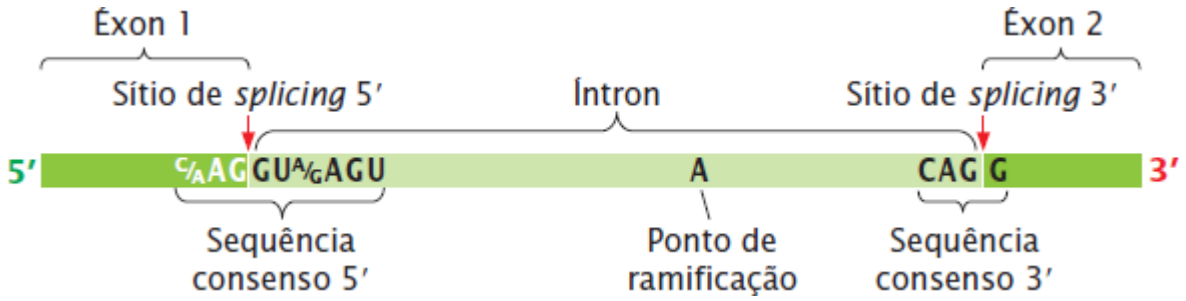


Figura 14.8 O *splicing* de pré-mRNA requer sequências consenso. Sequências consenso críticas são encontradas no sítio de *splicing* 5′ e no sítio de *splicing* 3′. Há uma sequência consenso fraca (não apresentada) no ponto de ramificação.

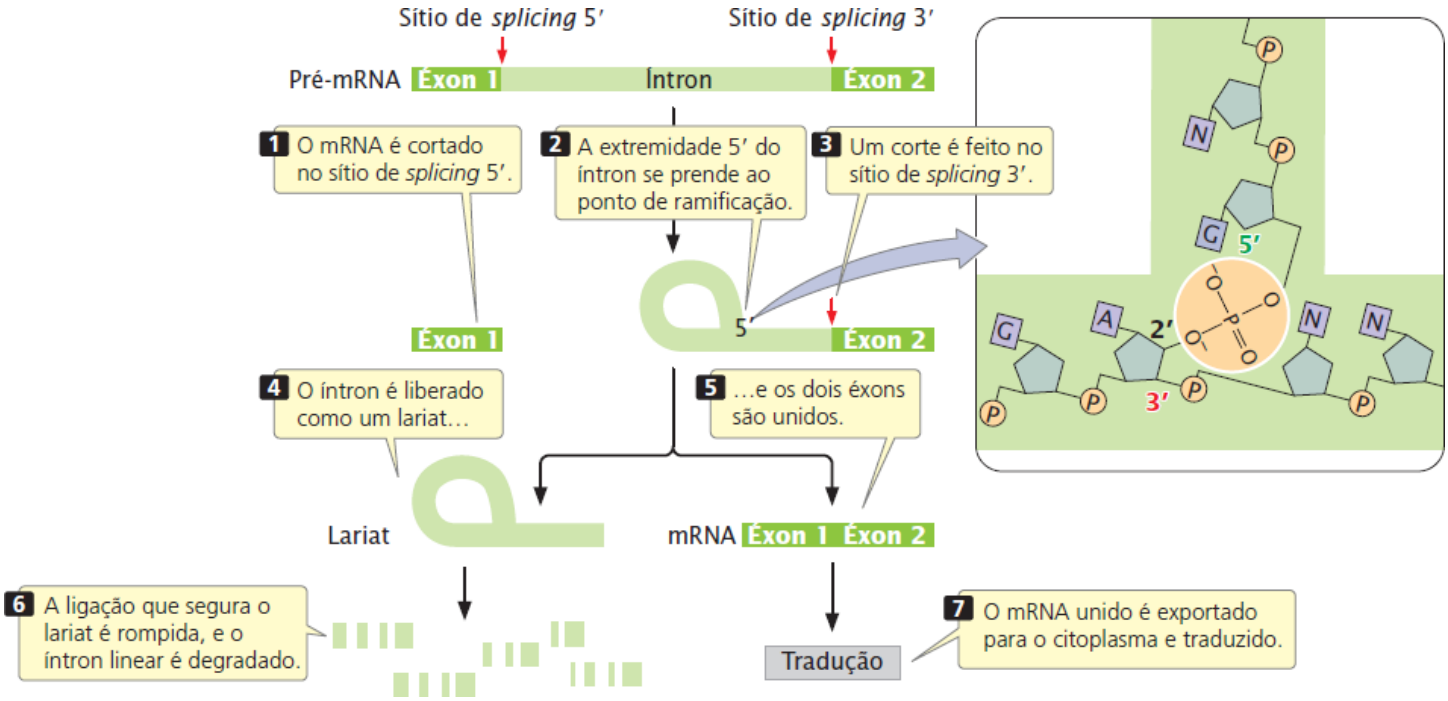


Figura 14.9 O *splicing* de íntrons nuclear requer um processo em duas etapas.

Na segunda etapa do *splicing* do RNA, um corte é feito no sítio de *splicing* 3′ e, simultaneamente, surge uma ligação covalente entre a extremidade do éxon 1 (unida) e a extremidade 5′ do éxons 2. O íntron é liberado como um lariat. Consequentemente, uma enzima que rompe a ramificação de lariat também rompe a ligação no ponto de ramificação, produzindo um íntron linear que é rapidamente degradado por enzimas nucleares. O mRNA maduro composto por éxons unidos é exportado para o citoplasma, onde é traduzido.

Essas reações de *splicing* ocorrem no spliceossomo, que é montado no pré-mRNA sequencialmente e faz as reações de *splicing* (**Figura 14.10**). Uma característica crítica do processo consiste nas várias interações entre o mRNA e os snRNAs e entre diferentes snRNAs. Essas interações dependem de pareamento de base complementar entre diferentes moléculas de RNA e trazem os componentes essenciais do pré-mRNA transcrito e o spliceossomo unidos, que tornam possível o *splicing*. As etapas catalíticas importantes no processo de *splicing* são feitas pelos snRNAs que compõem o spliceossomo.

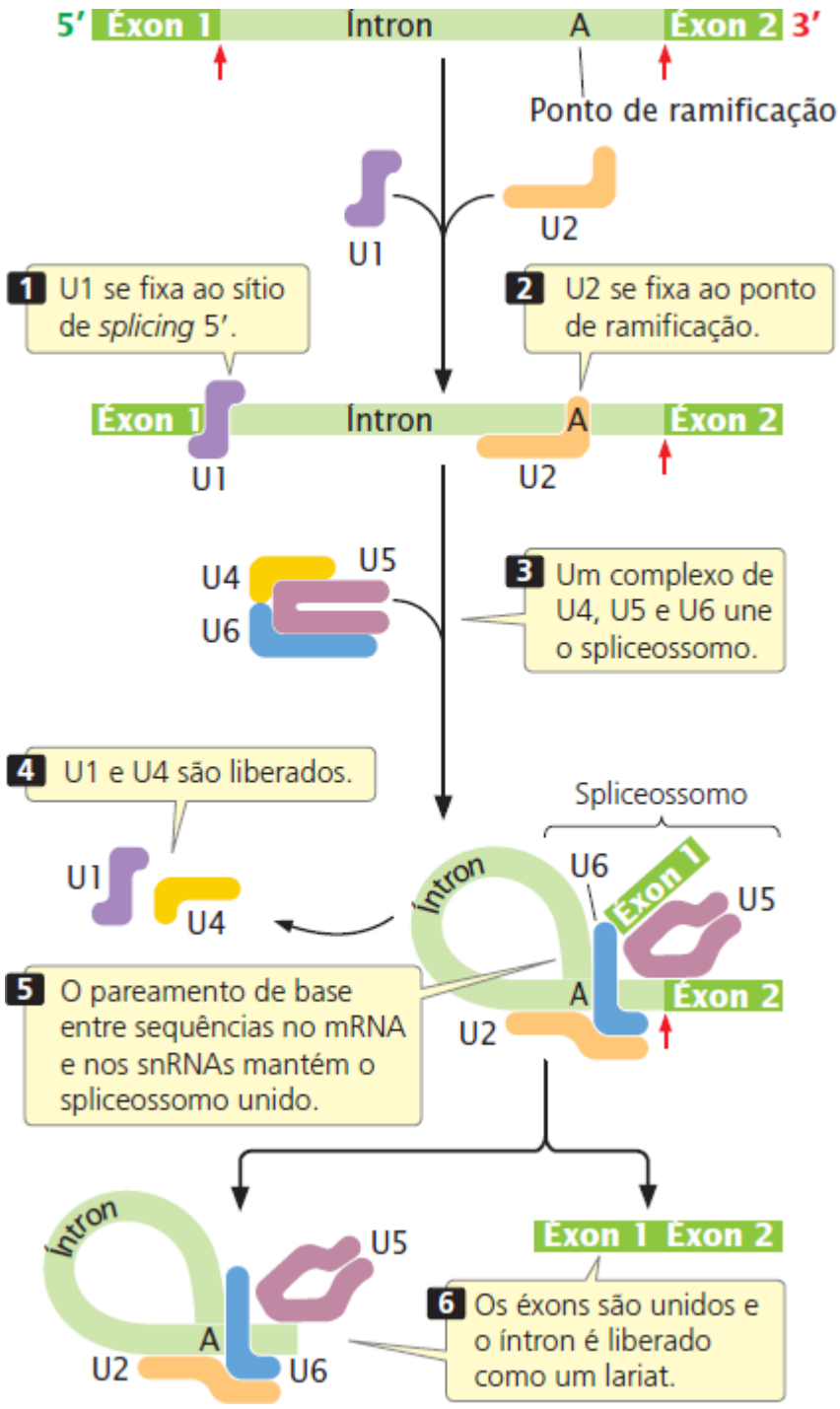


Figura 14.10 O *splicing* do RNA ocorre no spliceossomo, que, por sua vez, é montado sequencialmente.

Primeiro, snRNP U1 se prende ao sítio de *splicing* 5′ e então U2 se prende ao ponto de ramificação. Um complexo composto por U4, U5 e U6 (que formam um único snRNP) une o spliceossomo. Essa adição provoca uma mudança conformacional no spliceossomo, o íntron forma um laço e o sítio de *splicing* 5′ se aproxima do ponto de ramificação. As partículas U1 e U4 se dissociam do spliceossomo, com a subsequente formação de pares de base entre U6 e U2 e entre U6 e o sítio de *splicing* 5′. O sítio de *splicing* 5′, o sítio de *splicing* 3′ e o ponto de ramificação estão próximos, mantidos unidos pelo spliceossomo. As duas reações de transesterificação ocorrem, unindo os dois éxons e liberando o íntron como um lariat.

A maioria dos mRNAs é produzida a partir de uma única molécula de pré-mRNA a partir da qual os éxons são unidos. Entretanto, em alguns organismos (principalmente nos nematódeos e nos tripanossomos), os mRNAs podem ser produzidos ao unir sequências a partir de duas ou mais moléculas de RNA diferentes. Esse processo é chamado de ***trans-splicing*** ou recomposição trans.

Muitas doenças genéticas nos seres humanos surgem a partir de mutações que afetam o *splicing* do pré-mRNA; na verdade, cerca de 15% das substituições de uma base que resultam em doenças genéticas nos humanos alteram o *splicing* do pré-mRNA. Algumas dessas mutações interferem com a identificação dos sítios de *splicing* 5′ e 3′ normais. Outras criam novos sítios de *splicing*, como foi o caso com a mutação que causou a hemofilia na Família Real, discutida na introdução deste capítulo.

Resolva o Problema 24

O *splicing* do RNA, que ocorre no núcleo, tem de ser feito antes que o RNA possa se deslocar para o citoplasma. RNAs incompletamente unidos permanecem no núcleo até que o *splicing* esteja completo ou até que o pré-mRNA seja degradado. Imediatamente após o *splicing*, um grupo de proteínas chamadas de complexo de junção de éxon (EJC) é depositado em aproximadamente 20 nucleotídios *upstream* de cada junção éxon-éxon no mRNA. O EJC promove a exportação do mRNA do núcleo para o citoplasma.

Conceitos

O *splicing* de íntrons dos genes nucleares é um processo de duas etapas: (1) a extremidade 5′ do íntron é clivada e presa ao ponto de ramificação para formar um lariat e (2) a extremidade 3′ do íntron é clivada e as extremidades dos dois éxons são unidas. No processo, os éxons são unidos e o íntron interveniente é removido. Essas reações ocorrem dentro do spliceossomo.

Splicing menor. Alguns íntrons nos pré-mRNAs dos eucariotos multicelulares usam um processo diferente de remoção, chamado de *splicing* menor. Os íntrons que sofrem *splicing* menor apresentam diferentes sequências consenso no sítio de *splicing* 5′ e no ponto de ramificação e usam um spliceossomo menor, com um conjunto diferente de snRNAs. Cerca de 700 a 800 genes no genoma humano têm íntrons que sofrem *splicing* menor.

Íntrons auto-splicing. Alguns íntrons são do tipo *auto-splicing* ou autorremovíveis; eles têm a capacidade de se remover de uma molécula de RNA. Esses íntrons *auto-splicing* são divididos em duas categorias principais. Os íntrons do grupo I são encontrados em vários genes, incluindo alguns genes rRNA nos protistas, alguns genes mitocondriais nos fungos e até em alguns genes de bactérias e bacteriófagos. Embora o comprimento dos íntrons do grupo I varie muito, todos se dobram em uma estrutura secundária comum com nove hastes em alça (**Figura 14.11 A**), necessárias para o *splicing*.

Os íntrons do grupo II, encontrados nos genes das eubactérias, da archaea e das organelas eucarióticas, também têm a capacidade de se autorrecompor. Todos os íntrons do grupo II também se dobram em estruturas secundárias (**Figura 14.11 B**). O *splicing* dos íntrons do grupo II é feito por um mecanismo com algumas semelhanças em relação ao *splicing* mediado pelo spliceossomo nos genes nucleares, e o *splicing* gera uma estrutura lariat. Por causa dessas semelhanças, foi sugerido que os íntrons do grupo II e os íntrons pré-mRNA nucleares estejam próximos em termos evolutivos; talvez os íntrons nucleares tenham evoluído a partir dos íntrons *auto-splicing* do grupo II e, posteriormente, adotaram as proteínas e os snRNAs do spliceossomo para fazer a reação de *splicing*.

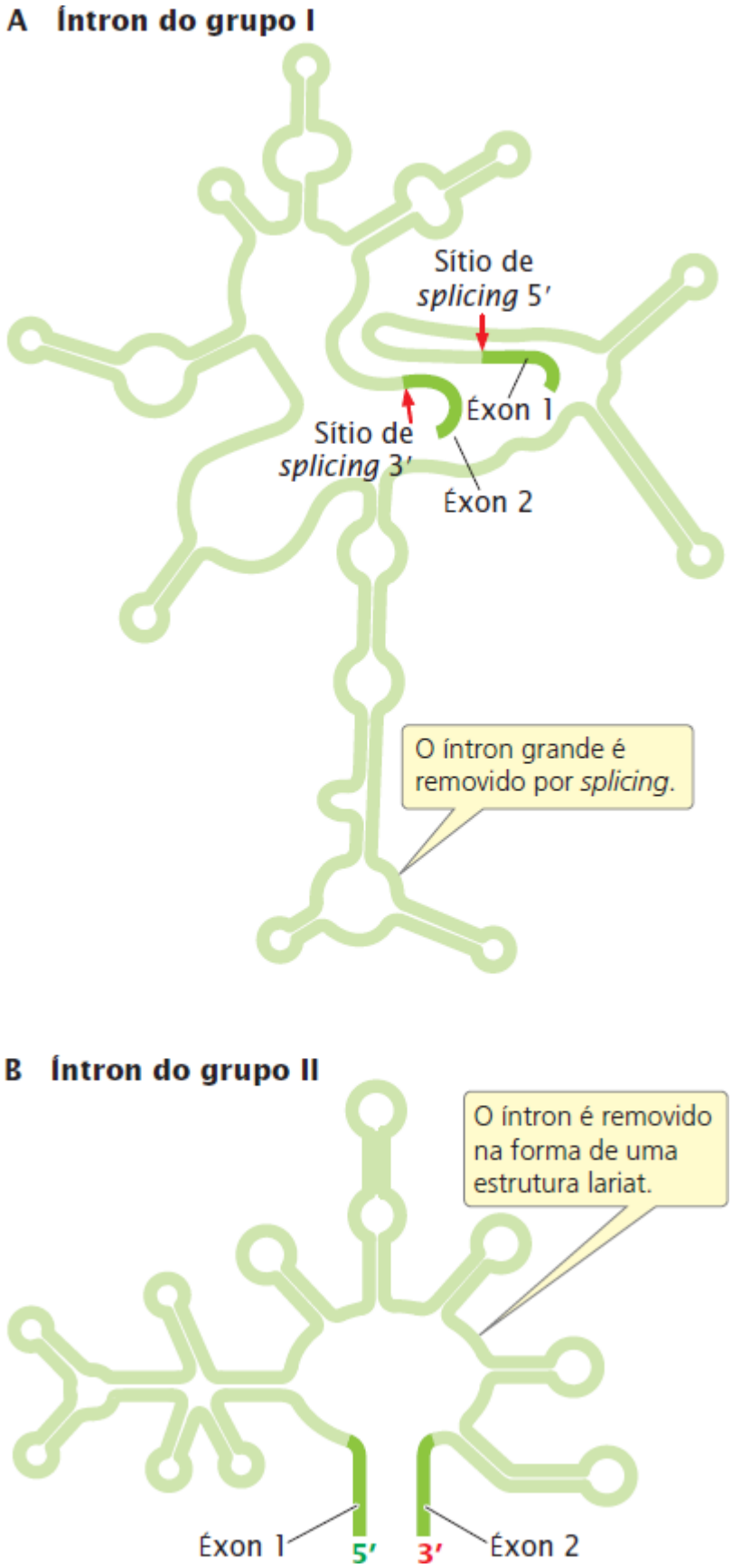


Figura 14.11 **Íntrons dos grupos I e II se dobram em estruturas secundárias características.**

Conceitos

Alguns íntrons são removidos pelo sistema de *splicing* menor. Outros são *auto-splicing* e compostos por dois tipos: íntrons do grupo I e íntrons do grupo II. Esses íntrons têm estruturas secundárias complexas que possibilitam catalisar sua excisão das moléculas de RNA sem a ajuda de enzimas ou de outras proteínas.

Vias de processamento alternativas

Um achado que complica a visão de um gene como uma sequência de nucleotídios que especifica a sequência de aminoácidos de uma proteína (como observado em *Conceito de gene revisitado*) é a existência de **vias de processamento alternativas**. Nessas vias, um único pré-mRNA é processado de diferentes maneiras para produzir tipos alternativos de mRNA, resultando na produção de diferentes proteínas a partir da mesma sequência de DNA.

Um tipo de processamento alternativo é o ***splicing* ou recomposição alternativa**, no qual o mesmo pré-mRNA pode ser unido em mais de uma maneira para gerar múltiplos mRNAs que são traduzidos em diferentes sequências de aminoácidos e diferentes proteínas (**Figura 14.12 A**). Outro tipo de processamento alternativo requer o uso de **sítios de clivagem 3′ múltiplos** (**Figura 14.12 B**), no qual dois ou mais sítios potenciais para clivagem e poliadenilação são encontrados no pré-mRNA. No exemplo da Figura 14.12 B, a clivagem no primeiro sítio produz um mRNA relativamente curto se comparado ao mRNA produzido por meio da ruptura do segundo sítio. O uso de um sítio de clivagem alternativo pode ou não produzir uma proteína diferente, dependendo se a posição do sítio estiver antes ou após o códon de término.

Ambos *splicing* alternativo e sítios de clivagem 3′ múltiplos podem existir no mesmo pré-mRNA transcrito. Um exemplo é observado no gene de mamíferos que codifica a calcitonina; esse gene tem seis éxons e cinco íntrons (**Figura 14.13 A**). O gene inteiro é transcrito em pré-mRNA (**Figura 14.13 B**). Existem dois sítios de clivagem 3′ possíveis. Nas células da glândula tireoide, a clivagem em 3′ e a poliadenilação ocorrem após o quarto éxon para produzir um mRNA maduro composto pelos éxons 1, 2, 3 e 4 (**Figura 14.13 C**). Esse mRNA é traduzido no hormônio calcitonina, produzido pela glândula tireoide e que regula os níveis de cálcio. Nas células do cérebro, o pré-mRNA *idêntico* é transcrito a partir do DNA, mas a clivagem e a poliadenilação ocorrem após o sexto éxon, gerando um transcrito inicial que inclui todos os seis éxons. Durante o *splicing*, o éxon 4 é removido, e então apenas os éxons 1, 2, 3, 5 e 6 são encontrados no mRNA maduro (**Figura 14.13 D**). Quando traduzido, esse mRNA produz uma proteína chamada peptídio relacionado com o gene da calcitonina (CGPR), com uma sequência de aminoácidos bem diferente da calcitonina. O CGPR provoca a dilatação dos vasos sanguíneos e pode atuar na transmissão da dor. Algumas pesquisas sugerem que ele esteja envolvido no desenvolvimento da enxaqueca. O *splicing* alternativo pode produzir diferentes combinações de éxons no mRNA, mas a ordem dos éxons não é alterada.

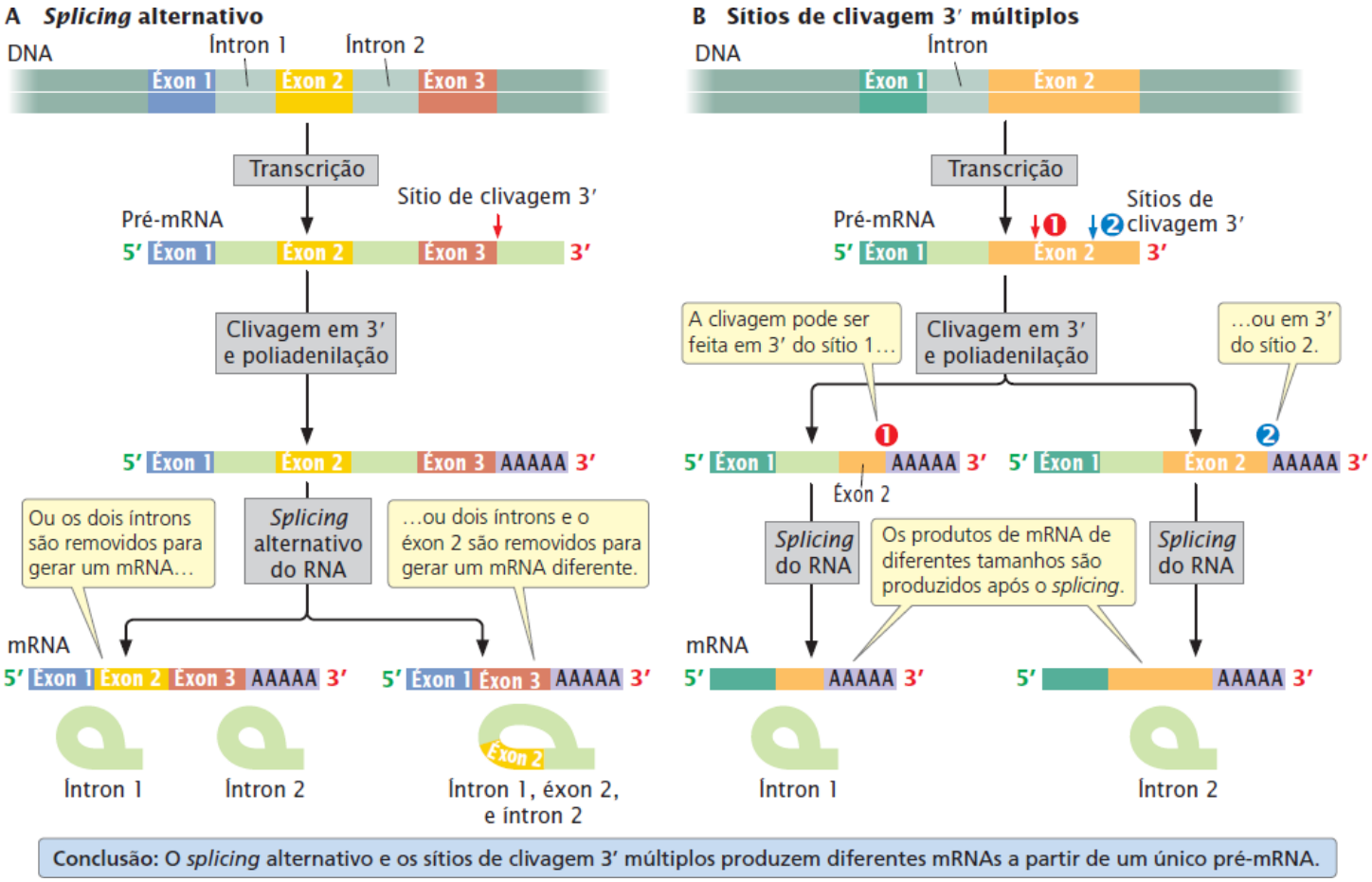


Figura 14.12 As células eucarióticas têm vias alternativas para o processamento do pré-mRNA. A. Graças ao *splicing* alternativo, o pré-mRNA pode ser recomposto de diferentes maneiras para produzir diferentes mRNAs. **B.** Com sítios de clivagem 3′ múltiplos, existem dois ou mais sítios potenciais para a clivagem e a poliadenilação; o uso de diferentes sítios produz mRNAs de diferentes tamanhos.

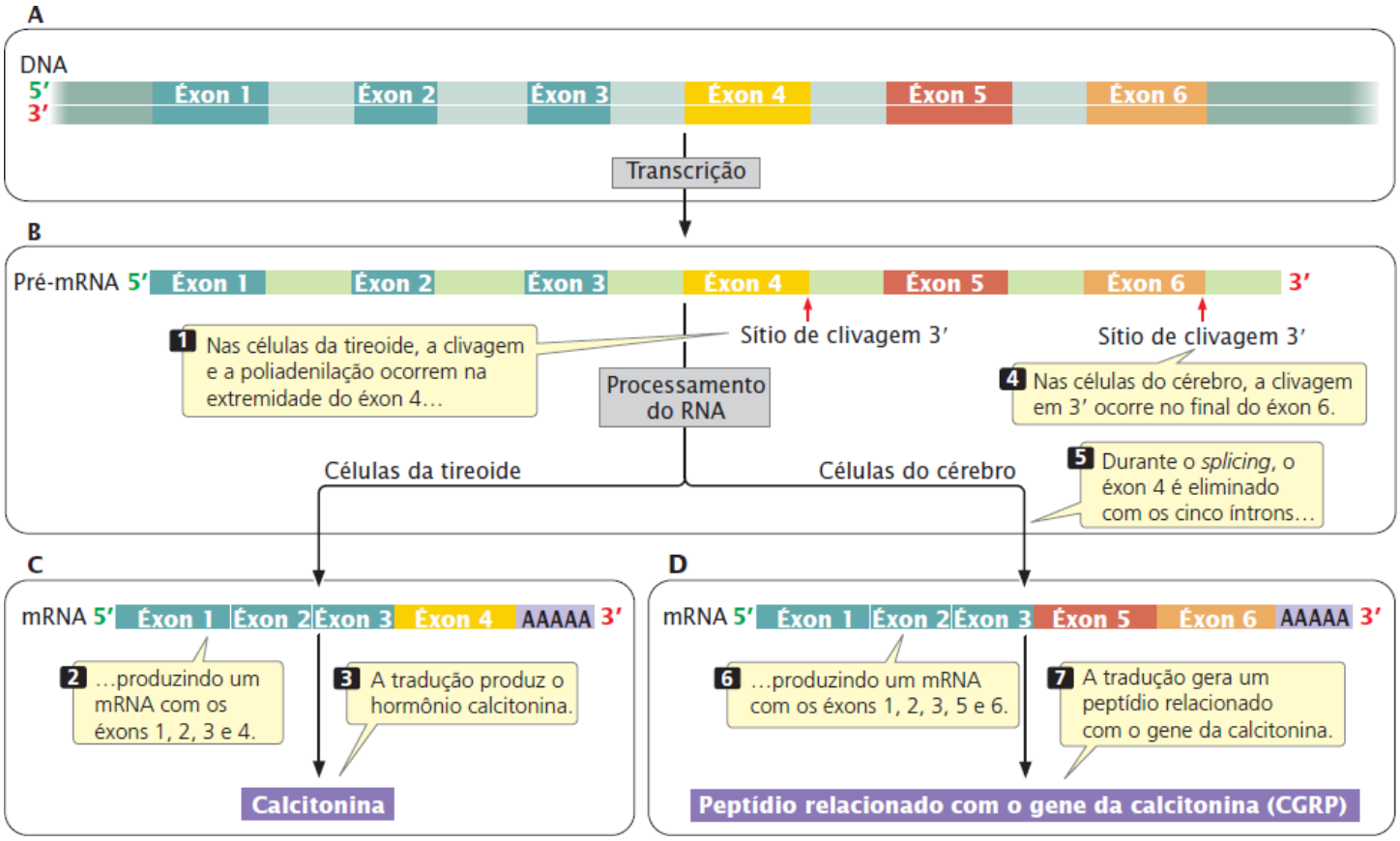


Figura 14.13 O pré-mRNA codificado pelo gene para calcitonina sofre processamento alternativo.

O processamento alternativo de pré-mRNAs é comum nos eucariotos multicelulares. Por exemplo, os pesquisadores estimam que mais de 90% de todos os genes humanos sofrem *splicing* alternativo. Muitas vezes a forma de *splicing* é diferente entre os tecidos humanos: os tecidos do cérebro e do fígado humanos têm mais *splicing* alternativo quando comparados com outros tecidos. Às vezes o *splicing* varia até de uma pessoa para outra. As diferentes vias de processamento contribuem para a regulação gênica, que será discutida no Capítulo 17.

O *splicing* alternativo participa na complexidade do organismo. O sequenciamento completo dos genomas de vários organismos (ver *Capítulo 20*) levou à conclusão de que o número de genes no organismo não está correlacionado com a complexidade do mesmo. Por exemplo, as moscas-da-fruta têm apenas 14.000 genes, enquanto o nematódeo mais simples anatomicamente tem 19.000 genes. A planta *Arabidopsis thaliana* tem cerca de 20.000 genes, quase tanto quanto os humanos. Se organismos anatomicamente simples têm tantos genes quanto os complexos, como a complexidade do desenvolvimento é codificada no genoma? Uma possível resposta é o processamento alternativo, que pode produzir múltiplas proteínas a partir de um único gene e é uma fonte importante de diversidade de proteínas nos vertebrados.

Pesquisa recente demonstrou que mesmo espécies muito próximas apresentam diferenças em como seus pré-mRNAs sofrem *splicing*, e o *splicing* alternativo é importante na especiação (ver *Capítulo 26*).

Conceitos

O *splicing* alternativo possibilita que os éxons sejam combinados de modo diferente para gerar mRNAs que codificam diferentes proteínas. Os sítios de clivagem 3′ alternativos tornam possível que o pré-mRNA seja clivado em diferentes sítios.

Checagem dos conceitos 5

Os sítios de clivagem 3′ alternativos resultam em:

a. Múltiplos genes de diferentes comprimentos.
b. Múltiplos pré-mRNAs de diferentes comprimentos.
c. Múltiplos mRNAs de diferentes comprimentos.
d. Todas as opções acima.

Edição do RNA

A hipótese de que todas as informações sobre a sequência de aminoácidos de uma proteína estão localizadas no DNA é contestada por um processo chamado **edição do RNA**. Nela, a sequência codificadora de uma molécula de mRNA é alterada após a transcrição, então a proteína tem uma sequência de aminoácidos diferente da sequência codificada pelo gene.

A edição do RNA foi detectada pela primeira vez em 1986, quando as sequências codificadoras de mRNAs foram comparadas com as sequências codificadoras de DNA a partir das quais elas foram transcritas. Em alguns genes nucleares nas células de mamíferos e em alguns genes mitocondriais nas células de plantas existem substituições em determinados nucleotídios do mRNA. Foi descoberta uma edição de RNA mais abrangente no mRNA para alguns genes mitocondriais nos parasitas tripanossomos (que causam a doença do sono na África). Em alguns mRNAs desses organismos, mais de 60% da sequência é determinada pela edição do RNA. Já foram observados diferentes tipos de edição de RNA nos mRNAs, tRNAs e rRNAs para uma ampla gama de organismos; os tipos incluem a inserção e a eliminação de nucleotídios e a conversão de uma base em outra.

Se a sequência modificada em uma molécula de RNA editada não provém de um molde de DNA, então como ela é especificada? Vários mecanismos podem fazer mudanças nas sequências de RNA. Em alguns casos, as moléculas chamadas de **RNAs guias** (gRNAs) exercem um papel crucial. Um gRNA tem sequências que são parcialmente complementares a segmentos de RNA pré-editado, e as duas moléculas sofrem pareamento de base nessas sequências (**Figura 14.14**). Após o mRNA ser ancorado ao gRNA, ele sofre clivagem, e nucleotídios são adicionados, eliminados ou alterados de acordo com o molde fornecido pelo gRNA. Em outros casos, as enzimas são responsáveis pela conversão de bases. Nos seres humanos, por exemplo, um gene é transcrito em mRNA que codifica um polipeptídio transportador de lipídio, chamado de apolipoproteína-B100, que tem 4.563 aminoácidos e é sintetizado nas células hepáticas. Uma forma truncada da proteína, chamada de apolipoproteína-B48 – com apenas 2.153 aminoácidos –, é sintetizada nas células do intestino por meio da edição do mRNA da apolipoproteína-B100. Nessa edição, uma enzima desamina uma base citosina, convertendo-a em uracila. Essa conversão muda um códon que especifica o aminoácido glutamina em um códon de parada que termina de maneira prematura a tradução, gerando em uma proteína encurtada. **Resolva o Problema 34**

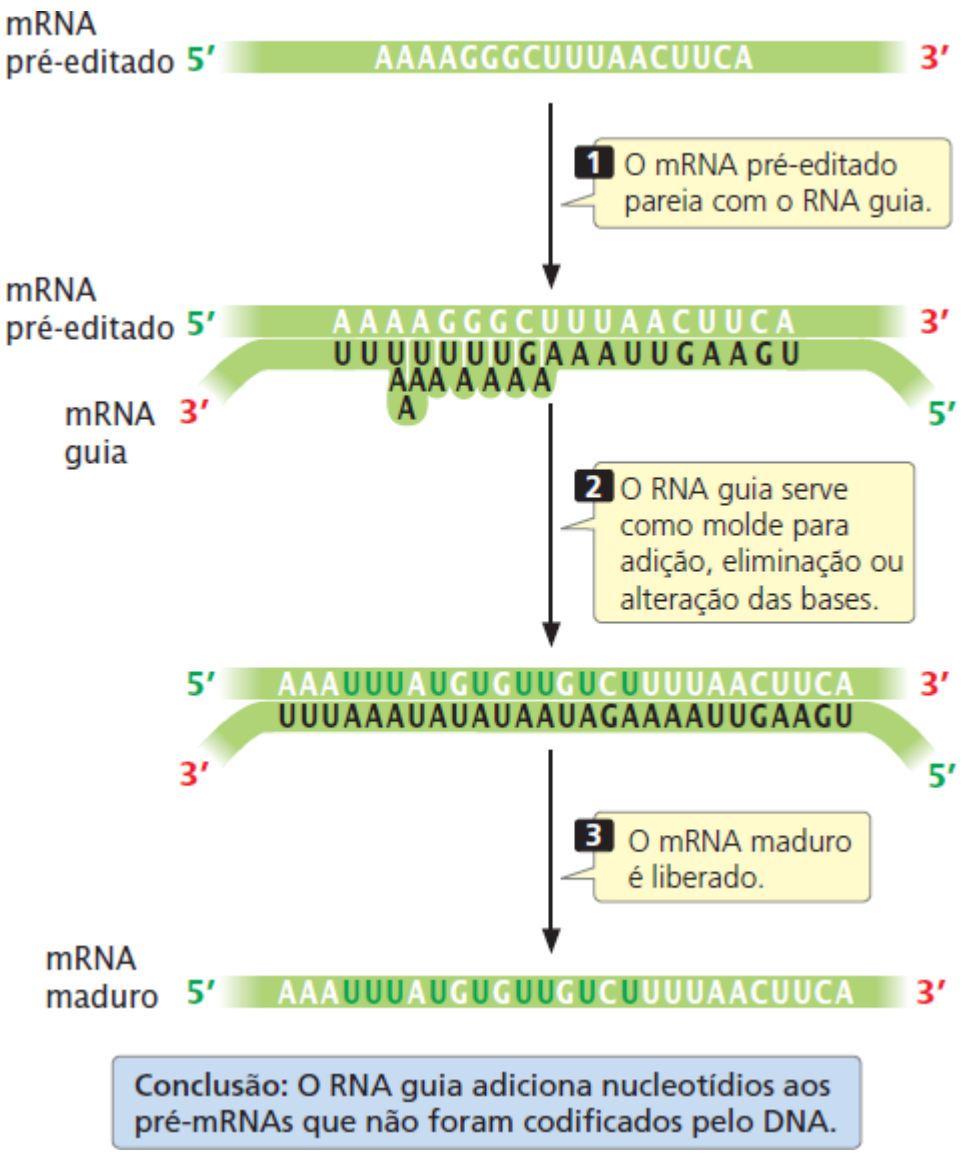


Figura 14.14 A edição do RNA é feita por RNAs guias. O mRNA guia tem sequências parcialmente complementares às do mRNA pré-editado e pareia com este. Após o pareamento, o mRNA sofre clivagem e novos nucleotídios são adicionados, com sequências no gRNA servindo como molde. As extremidades do mRNA são unidas.

Conceitos

Nucleotídios no interior de pré-mRNA podem ser alterados, adicionados ou eliminados pela edição do RNA. A sequência de aminoácidos produzida pelo mRNA editado não é a mesma codificada pelo DNA.

Checagem dos conceitos 6

O que especifica a sequência modificada de nucleotídios encontrados em uma molécula de RNA editada?

Conceitos conectantes

Estrutura do gene eucariótico e processamento do pré-mRNA

Os Capítulos 13 e 14 descrevem vários componentes diferentes de genes e moléculas de RNA, incluindo promotores, regiões 5′ não traduzidas, sequências codificadoras, íntrons, regiões 3′ não traduzidas, caudas poli(A) e estruturas cap. Vejamos como alguns desses componentes são combinados para criar um típico gene eucariótico e como um mRNA maduro é produzido a partir desses componentes.

O promotor, que fica *upstream* do sítio de início da transcrição, é necessário para que a transcrição ocorra, mas ele não é transcrito quando os genes codificadores de proteína são transcritos pela RNA polimerase II (**Figura 14.15 A**). Mais distante *upstream* ou *downstream* do sítio de início estão os acentuadores, sequências de DNA que também regulam a transcrição.

Nesse processo, todos os nucleotídios entre o sítio de início de transcrição e o sítio de término são transcritos em pré-

mRNA, incluindo éxons, íntrons e uma longa extremidade 3′ que é posteriormente clivada a partir do transcrito (**Figura 14.15 B**). Observe que a extremidade 5′ do primeiro éxon contém a sequência que codifica a região 5′ não traduzida, e que a extremidade 3′ do último éxon contém a sequência que codifica a região 3′ não traduzida.

O pré-mRNA é então processado para gerar um mRNA maduro. A primeira etapa nesse processamento é a adição de um cap à extremidade 5′ do pré-mRNA (**Figura 14.15 C**). A seguir, a extremidade 3′ é clivada em um sítio *downstream* da sequência consenso AAUAAA no último éxon (**Figura 14.15 D**). Imediatamente após a clivagem, uma cauda poli(A) é adicionada à extremidade 3′ (**Figura 14.15 E**). Finalmente, os íntrons são removidos para gerar o mRNA maduro (**Figura 14.15 F**). Agora o mRNA tem as regiões 5′ e 3′ não traduzidas, que não são traduzidas em aminoácidos, e os nucleotídios que carreiam as sequências codificadoras de proteína.

A **Figura 14.16** apresenta a sequência de nucleotídios de um pequeno gene (o gene da interleucina 2 humana) com esses componentes identificados.

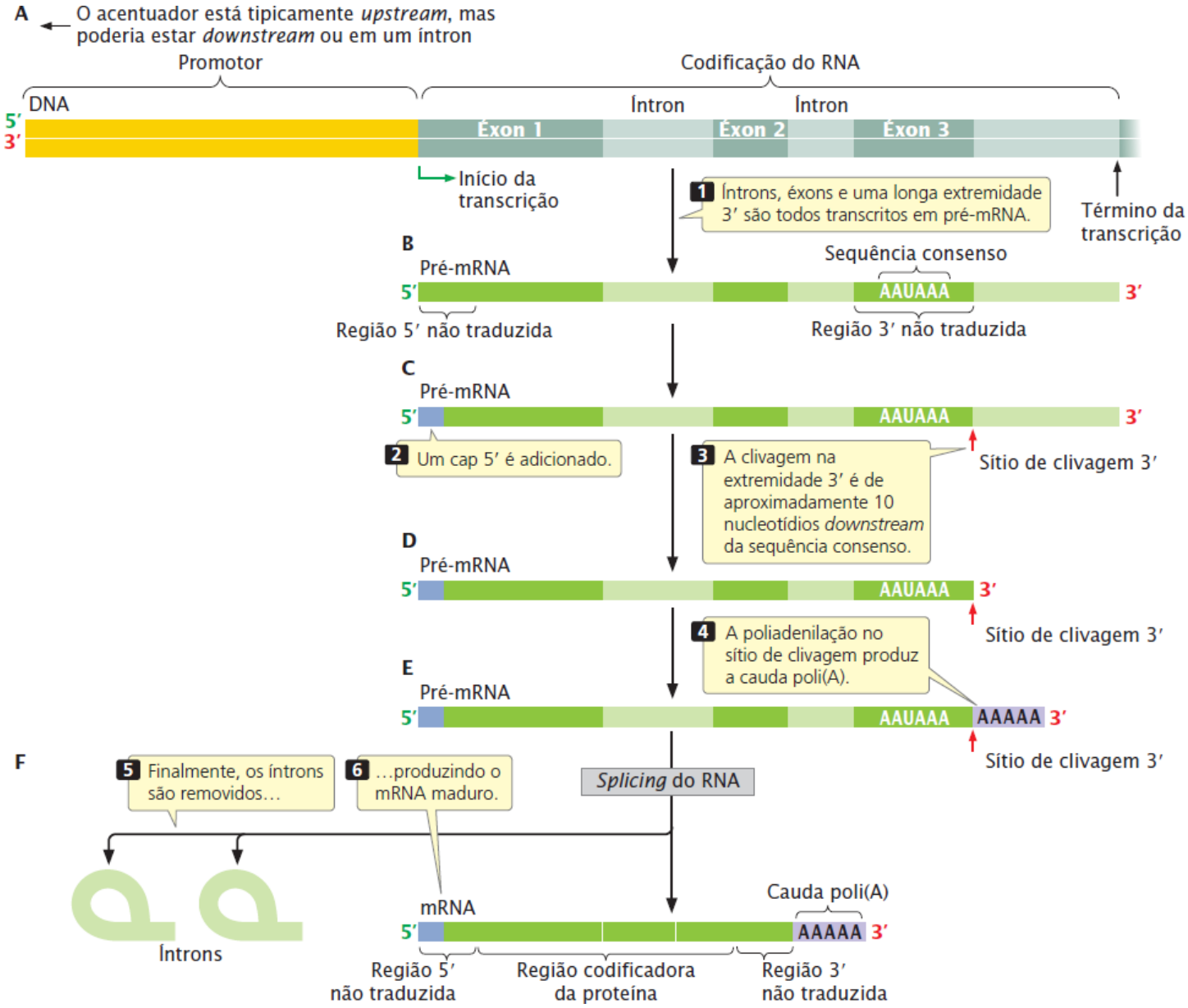


Figura 14.15 O mRNA eucariótico maduro é produzido quando o pré-mRNA é transcrito e sofre vários tipos de processamento.

Figura 14.16 **Esta representação da sequência de nucleotídios do gene para a interleucina 2 humana inclui TATA box, o sítio de início de transcrição, códons de início e parada, íntrons, éxons, sequência consenso poli(A) e o sítio de clivagem 3′.**

14.3 Os RNAs transportadores, ligados aos aminoácidos, são modificados após a transcrição nas células bacterianas e eucarióticas

Em 1956, Francis Crick propôs a ideia de uma molécula que transporta aminoácidos para o ribossomo e que interage com os códons no mRNA, colocando os aminoácidos em sua ordem correta na síntese de proteínas. Em 1963, a existência dessa molécula adaptadora, chamada de RNA transportador, foi confirmada. O RNA transportador (tRNA) serve como elo entre o código genético do mRNA e os aminoácidos que compõem uma proteína. Cada tRNA se liga a um aminoácido específico e o carreia até o ribossomo, onde o tRNA adiciona seu aminoácido à cadeia polipeptídica crescente na posição especificada pelas instruções genéticas contidas no mRNA. Vamos examinar o mecanismo desse processo com mais detalhes no Capítulo 15.

Cada tRNA consegue se ligar a apenas um tipo de aminoácido. O complexo de tRNA mais seu aminoácido podem ser escritos no modo abreviado ao adicionar um superscrito de três letras, representando o aminoácido no tRNA designado. Por exemplo, um tRNA que se prende ao aminoácido alanina é escrito como $tRNA^{Ala}$. Como são encontrados 20 tipos diferentes de aminoácidos nas proteínas, existem pelo menos 20 tipos diferentes de tRNA. Na verdade, a maioria dos organismos tem pelo menos 30 a 40 tipos diferentes de tRNA, cada um codificado por um gene diferente (ou, em alguns casos, múltiplas cópias de um gene) no DNA.

Estrutura do RNA transportador

Uma característica particular do tRNA é a ocorrência de **bases modificadas** raras. Todos os RNAs têm quatro bases padrão (adenina, citosina, guanina e uracila) especificadas pelo DNA, mas os tRNAs têm bases adicionais, incluindo ribotimina, pseudouridina (ocasionalmente encontrada no snRNAs e no rRNA) e diversas outras. As estruturas de duas dessas bases modificadas são apresentadas na **Figura 14.17**.

Se existem apenas quatro bases no DNA e todas as moléculas de RNA são transcritas a partir do DNA, como os tRNAs adquirem essas bases adicionais? As bases modificadas surgem a partir de mudanças químicas ocorridas nas quatros bases padrão após a transcrição. Essas mudanças são feitas por **enzimas modificadoras de tRNA** especiais. Por exemplo, a adição de um grupo metila à uracila cria a base modificada ribotimidina.

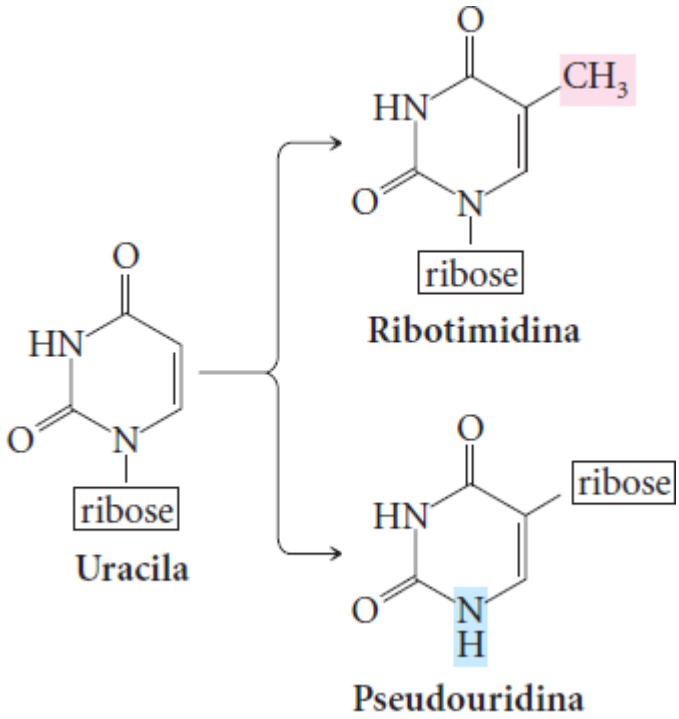


Figura 14.17 Duas das bases modificadas encontradas nos tRNAs. Todas as bases modificadas nos tRNAs são produzidas por alteração química das quatro bases RNA padrão.

As estruturas de todos os tRNAs são semelhantes, uma característica crítica para a função do tRNA. A maioria dos tRNAs tem entre 74 e 95 nucleotídios, alguns dos quais são complementares com outros e formam pontes de hidrogênio intramoleculares. Como resultado, cada tRNA tem uma **estrutura de folha de trevo** (**Figura 14.18**). A folha de trevo tem quatro braços principais. Se começarmos do topo e seguirmos no sentido horário ao redor do tRNA apresentado à direita na Figura 14.18, os quatro braços principais são o braço aceptor, o braço TψC, o braço anticódon e o braço DHU. Três desses braços (o TψC, o anticódon e o DHU) são compostos por uma haste e uma alça. A haste é formada pelo pareamento de nucleotídios complementares, e a alça fica no término da haste, onde não existe pareamento de nucleotídio.

Em vez de ter uma alça, o braço aceptor inclui as extremidades 5′ e 3′ da molécula de tRNA. Todos os tRNAs têm a mesma sequência (CCA) na extremidade 3′, onde o aminoácido se fixa ao tRNA; então, evidentemente, esta sequência não é responsável por especificar qual aminoácido se fixará ao tRNA.

O braço TψC é composto por bases de três nucleotídios em sua alça: timina (T), pseudouridina (ψ) e citosina (C). O braço anticódon fica na parte inferior do tRNA. Três nucleotídios na extremidade deste braço compõem o **anticódon**, que pareia com o códon correspondente no mRNA, para garantir que os aminoácidos se liguem na ordem correta. O braço DHU é chamado assim porque tem a base modificada di-hidrouridina.

Embora cada molécula de tRNA se dobre formando uma folha de trevo em função do pareamento complementar das bases, o trevo não é a estrutura tridimensional (terciária) dos tRNAs encontrados na célula. Os resultados de estudos de cristalografia pelos raios X mostraram que a folha de trevo se dobra sobre si mesma em uma estrutura com formato em L, como ilustrado nos modelos de preenchimento de espaço e de fita na Figura 14.18. Observe que a haste aceptora está em uma extremidade da estrutura terciária e o anticódon está em outra.

Estrutura e processamento do gene do RNA transportador

O gene que produz tRNAs pode estar em grupos ou disperso no genoma. Na *E. coli*, os genes para alguns tRNAs são encontrados em uma cópia única, enquanto os genes para outros tRNAs são vistos em diversas cópias; as células eucarióticas, em geral, apresentam muitas cópias de cada gene de tRNA. Todas as moléculas de tRNA – tanto nas células bacterianas quanto nas eucarióticas – sofrem processamento após a transcrição.

Na *E. coli*, muitos tRNAs são transcritos juntos, como um grande tRNA precursor, que é então cortado em pedaços, cada um contendo um único tRNA. Nucleotídios adicionais podem então ser removidos um por vez nas extremidades 5′ e 3′ do tRNA, em um processo conhecido como aparo. As enzimas modificadoras de base, por sua vez, podem alterar algumas das bases padrão em bases modificadas (**Figura 14.19**). Em alguns procariotos, as sequências CCA encontradas nas extremidades 3′ dos tRNA são codificadas no gene do tRNA e transcritas em tRNA; em outros procariotos e eucariotos, essas sequências são adicionadas por uma enzima especial que adiciona os nucleotídios sem o uso de molde.

Não existe uma via genérica de processamento para todos os tRNAs: diferentes tRNAs são processados de diferentes

maneiras. Os tRNAs eucarióticos são processados de um modo semelhante aos tRNAs bacterianos: a maioria é transcrita como precursores maiores que são então clivados, aparados e modificados para produzir tRNAs maduros.

Alguns genes de tRNAs eucarióticos e de archaea têm íntrons de tamanho variável que devem ser removidos no processamento. Por exemplo, cerca de 40 dos 400 genes de tRNA na levedura têm um único íntron que é sempre encontrado adjacente ao sítio 3′ do anticódon. Os íntrons de tRNA são mais curtos que os encontrados no pré-mRNA e não têm sequências consenso nas junções íntron-éxon dos pré-mRNAs. O processo de *splicing* para os genes tRNA é bem diferente das reações mediadas pelo spliceossomo, que removem íntrons dos genes codificadores de proteínas.

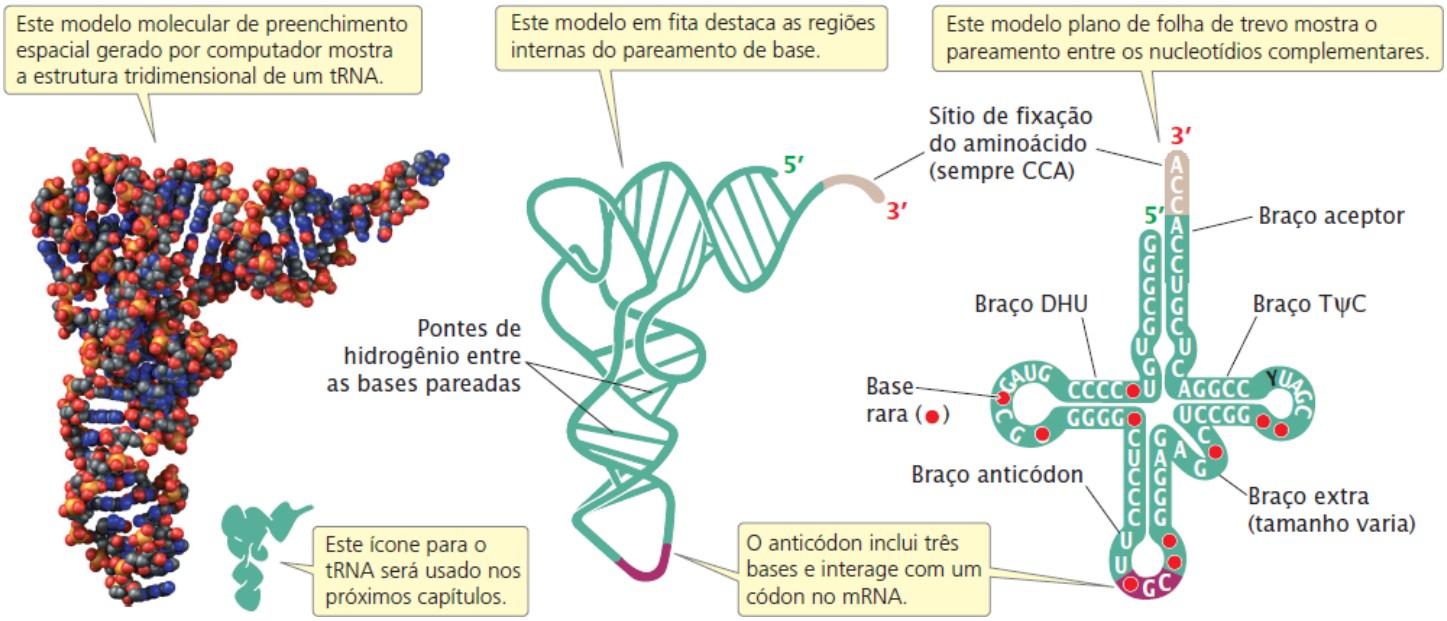


Figura 14.18 Todos os tRNAs têm uma estrutura secundária comum, a em folha de trevo. A sequência de bases no molde plano é para $tRNA^{Ala}$.

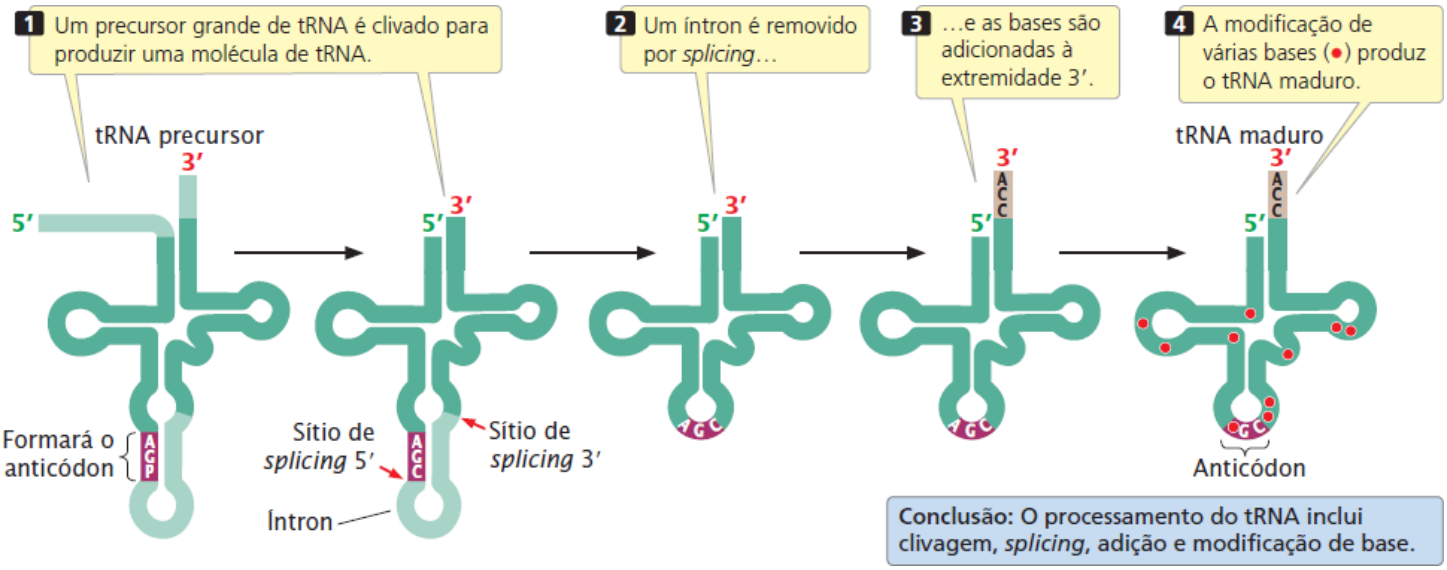


Figura 14.19 Os RNAs transportadores são processados nas células bacterianas e eucarióticas. Diferentes tRNAs são modificados de diferentes maneiras. Um exemplo é descrito aqui.

Conceitos

Todos os tRNAs têm tamanho semelhante e uma estrutura secundária comum conhecida como folha de trevo. Os RNAs transportadores têm bases modificadas que são substancialmente processadas após a transcrição tanto nas células bacterianas quanto nas eucarióticas.

Checagem dos conceitos 7

Como bases raras são incorporadas aos tRNAs?

a. Codificadas por RNAs guias.

b. Por mudanças químicas em uma das bases padrão.
c. Codificadas por bases raras no DNA.
d. Codificadas por sequências em íntrons.

14.4 O RNA ribossômico, um componente do ribossomo, também é processado após a transcrição

Nos ribossomos, as instruções genéticas contidas no mRNA são traduzidas em sequências de aminoácidos de polipeptídios. Assim, os ribossomos são uma parte integral na transferência de informação genética do genótipo para o fenótipo. Examinaremos o papel dos ribossomos no processo de tradução no Capítulo 15. Aqui, vamos analisar a estrutura do ribossomo e examinar como eles são processados antes de se tornarem funcionais.

A estrutura do ribossomo

O ribossomo é um dos complexos moleculares mais abundantes na célula: uma única célula bacteriana pode ter até 20.000 ribossomos, e as células eucarióticas podem ter ainda mais. Tipicamente, os ribossomos têm cerca de 80% do RNA total da célula. São estruturas complexas, cada uma com mais de 50 proteínas e moléculas de RNA diferentes (**Quadro 14.3**). Um ribossomo funcional apresenta duas subunidades, uma **subunidade ribossômica maior** e uma **subnidade ribossômica menor**, cada uma com um ou mais pedaços de RNA e várias proteínas. Os tamanhos dos ribossomos e de seus componentes RNA são dados em unidades Svedberg (S) (uma medida da velocidade de sedimentação de um objeto em um campo centrífugo). É importante observar que as unidades S não são aditivas; combinar uma estrutura 10S e uma estrutura 20S não produz necessariamente uma estrutura 30S, porque a velocidade de sedimentação é afetada pela estrutura tridimensional, assim como pela massa. A estrutura tridimensional do ribossomo bacteriano foi elucidada com muitos detalhes por meio da cristalografia por raios X. O Capítulo 15 discute com mais detalhes a estrutura do ribossomo.

Estrutura e processamento do gene do RNA ribossômico

Os genes para rRNA, assim como os genes para tRNA, podem ser encontrados em múltiplas cópias, e os números variam entre as espécies (**Quadro 14.4**); todas as cópias de rRNA em uma espécie são idênticas ou quase idênticas. Nas bactérias, os genes do rRNA estão dispersos, mas nas células eucarióticas eles estão agrupados, com os genes dispostos em *tandem*, um após o outro.

Quadro 14.3 Composição dos ribossomos nas células bacterianas e eucarióticas.

Tipo de célula	Tamanho do ribossomo	Subunidade	Componente do rRNA	Proteínas
Bacteriano	70S	Maior (50S)	23S (2.900 nucleotídios), 5S (120 nucleotídios)	31
		Menor (30S)	16S (1.500 nucleotídios)	21
Eucariótico	80S	Maior (60S)	28S (4.700 nucleotídios), 5,8S (160 nucleotídios), 5S (120 nucleotídios)	49
		Menor (40S)	18S (1.900 nucleotídios)	33

Nota: A letra “S” representa a “unidade de Svedberg”.

Quadro 14.4 Número de genes de rRNA em diferentes organismos.

Espécies	Número de cópias de genes de rRNA por genoma
Escherichia coli	7
Levedura	100 a 200
Humano	280

Sapo	450

As células eucarióticas apresentam dois tipos de genes de rRNA: um gene maior que codifica 18S rRNA, 28S rRNA e 5,8S rRNA, e um gene menor, que codifica o 5S rRNA. Todos os três rRNAs bacterianos (23S rRNA, 16S rRNA e 5S rRNA) são codificados por um único tipo de gene.

O RNA ribossômico é processado nas células bacterianas e eucarióticas. Na *E. coli*, cada gene de RNA é transcrito em um precursor 30S rRNA (**Figura 14.20 A**). Esse precursor 30S é metilado em vários locais, e então é clivado e aparado para produzir 16S rRNA, 23S rRNA e 5S rRNA, junto com um ou mais tRNAs. Uma série de enzimas é responsável pela clivagem, pela metilação e pelo aparo.

Os rRNAs eucarióticos sofrem processamento semelhante (**Figura 14.20 B**). Pequenos RNAs nucleolares (snoRNAs) ajudam a romper e modificar os rRNAs eucarióticos e a montar os rRNAs processados em ribossomos maduros. Como os snRNAs participam no *splicing* do pré-mRNA, os snoRNAs associam-se com proteínas para formar partículas de ribonucleoproteínas (snoRNPs). Os snoRNAs têm grande complementaridade com as sequências de rRNA, onde ocorre a modificação. É interessante notar que alguns snoRNAs são codificados por sequências nos íntrons de outros genes codificadores de proteínas. O processamento de rRNA e a montagem do ribossomo nos eucariotos ocorrem no nucléolo.

Conceitos

Um ribossomo é uma organela complexa composta por várias moléculas de rRNA e muitas proteínas. Cada ribossomo funcional tem uma subunidade maior e uma menor. Os RNAs ribossômicos tanto nas células bacterianas quanto nas eucarióticas são modificados após a transcrição. Nos eucariotos, o processamento do rRNA é feito por pequenos RNAs nucleolares.

Checagem dos conceitos 8

Que tipos de mudanças ocorrem no processamento do rRNA?

a. Metilação de bases.
b. Clivagem de um precursor maior.
c. Nucleotídios são aparados a partir das extremidades dos rRNAs.
d. Todas as opções acima.

14.5 Pequenas moléculas de RNA participam em várias funções

Algumas evidências sugerem que o primeiro material genético foi o RNA, e nos primórdios do planeta a vida era dominada por moléculas de RNA (ver *Capítulo 13*). Esse período de tempo, quando o RNA dominou os processos essenciais da vida, foi chamado de “mundo primitivo do RNA”. A percepção comum é que esse mundo de RNA “morreu” bilhões de anos atrás, quando muitas funções do RNA foram substituídas por moléculas de DNA mais estáveis e catalisadores proteicos mais eficientes. Entretanto, nos últimos 10 anos foram descobertas numerosas moléculas de RNA pequenas (a maioria com 20 a 30 nucleotídios de comprimento) que influenciam substancialmente muitos processos biológicos básicos, incluindo a formação da estrutura da cromatina, a transcrição e a tradução. Essas pequenas moléculas de RNA exercem papéis importantes na expressão gênica, no desenvolvimento, no câncer e na defesa contra DNA estranho. Elas também estão sendo aproveitadas pelos pesquisadores para estudar a função dos genes e tratar as doenças genéticas. A descoberta de pequenas moléculas de RNA influenciou significativamente nossa compreensão de como os genes são regulados e da importância das sequências de DNA que não codificam proteínas. Esses novos achados demonstram que ainda vivemos em um mundo de RNA.

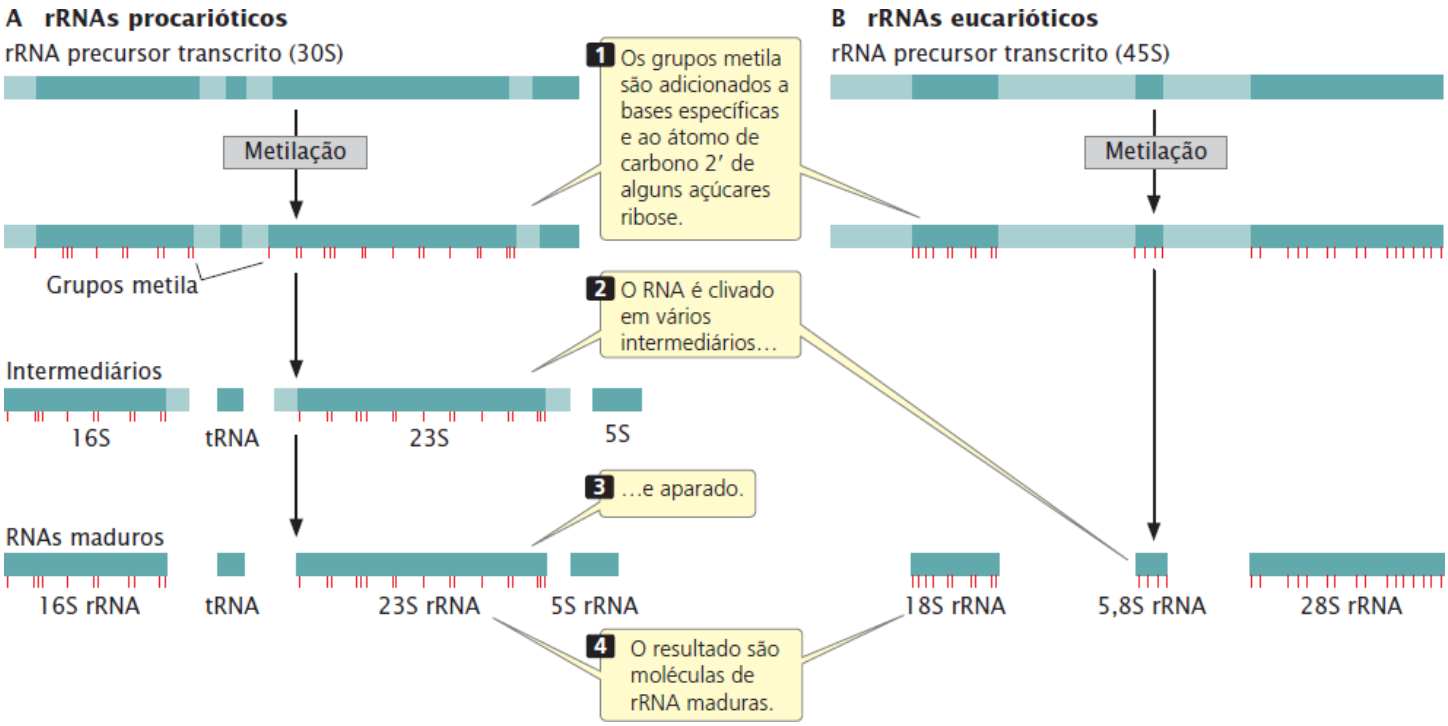


Figura 14.20 O RNA ribossômico é processado após a transcrição. O rRNA procariótico (**A**) e o rRNA eucariótico (**B**) são produzidos a partir de RNAs precursores transcritos que são metilados, clivados e processados para produzir rRNAs maduros. O 5S rRNA eucariótico é transcrito separadamente a partir de um gene diferente.

Nesta seção examinaremos a interferência do RNA (que levou à descoberta de pequenos RNAs), diferentes tipos de pequenos RNAs e como os miRNAs são processados. No Capítulo 17, vamos observar o papel dos pequenos RNAs no controle da expressão gênica; no Capítulo 19, veremos como os pequenos RNAs são usados como ferramentas importantes na biotecnologia.

Interferência por RNA

Em 1998, Andrew Fire, Craig Mello e seus colegas observaram um estranho fenômeno. Eles estavam inibindo a expressão de genes no nematódeo *Caenorhabditis elegans* ao inserir moléculas de RNA de fita única que eram complementares a uma sequência de gene de DNA. Chamado de RNA antissenso, tais moléculas são conhecidas por inibir a expressão gênica ao se ligar a sequências de mRNA inibindo a tradução. Fire, Mello e colaboradores descobriram que o silenciamento mais potente de gene era acionado quando o RNA de fita dupla era injetado nos animais. Tal descoberta era confusa, porque não era conhecido nenhum mecanismo pelo qual o RNA de fita dupla pudesse inibir a tradução. Foi descoberto que diversos outros tipos de silenciadores de gene eram acionados por RNA de fita dupla. Esses estudos iniciais levaram à descoberta de pequenas moléculas de RNA que eram importantes no silenciamento gênico.

Pesquisa posterior revelou uma ordem impressionante de pequenas moléculas de RNA com importantes funções celulares nos eucariotos, que agora incluem pelo menos três classes principais: pequenos RNAs de interferência (siRNAs), microRNAs (miRNAs) e RNAs que interagem com Piwi (piRNAs), dependendo de sua origem e modo de função. Esses pequenos RNAs são encontrados em muitos eucariotos e são responsáveis por várias funções diferentes, incluindo regulação da expressão gênica, defesa contra vírus, supressão dos transpósons e modificação da estrutura da cromatina. Um grupo análogo de pequenos RNAs com funções silenciadoras – chamados de CRISPR RNAs (crRNAs) – foram detectados nos procariotos. Fire e Mello receberam o prêmio Nobel em fisiologia ou medicina em 2006 por sua descoberta da interferência por RNA.

A **interferência por RNA (RNAi)** é um mecanismo poderoso e preciso empregado pelas células eucarióticas para limitar a invasão de genes estranhos (oriundos de vírus e transpósons) e para censurar a expressão de seus próprios genes. A interferência por RNA é acionada por moléculas de RNA de fita dupla, que podem surgir de várias maneiras (**Figura 14.21**): pela transcrição de repetições invertidas em uma molécula de RNA que então pareia a base com si mesma para formar RNA de fita dupla, pela transcrição simultânea de duas moléculas de RNA diferentes que são complementares uma com a outra e pareiam formando o RNA de fita dupla, ou pela infecção por vírus que criam o RNA de fita dupla. Essas moléculas de RNA de fita dupla são picotadas por uma enzima apropriadamente chamada Dicer, gerando moléculas de RNA minúsculas que são desenroladas para produzir siRNAs e miRNAs (ver Figura 14.21).

Alguns geneticistas especulam que a interferência por RNA evoluiu como um mecanismo de defesa contra os vírus de RNA e elementos transponíveis que se deslocam pelos intermediários do RNA (ver *Capítulo 18*); de fato, alguns pesquisadores chamam o RNAi de sistema imunológico do genoma. Entretanto, a interferência por RNA também é responsável por regular

vários processos genéticos e de desenvolvimento fundamentais, incluindo mudanças na estrutura da cromatina, tradução, destino, proliferação e morte celulares. Os geneticistas também usam o maquinário do RNAi como uma ferramenta efetiva para bloquear a expressão de genes específicos (ver *Capítulo 19*).

Pequenos RNAs de interferência e microRNAs

Duas classes abundantes de moléculas de RNA que atuam no RNA de interferência nos eucariotos são pequenos RNA de interferência e microRNAs. Embora esses dois tipos de RNA tenham diferentes origens (**Quadro 14.5**, ver Figura 14.21), eles apresentam várias características em comum, e suas funções se sobrepõem. Ambos têm 22 nucleotídios. Os pequenos RNAs de interferência surgem a partir da clivagem de mRNAs, RNA transpósons e vírus com RNA. Alguns miRNAs são clivados a partir de moléculas de RNA transcritas a partir de sequências que codificam apenas o miRNA, mas outras são codificadas nos íntrons e éxons dos mRNAs. Cada miRNA é clivado de um precursor RNA de fita única que forma pequenos grampos, enquanto múltiplos siRNAs são produzidos a partir da clivagem de um duplex de RNA composto por duas moléculas de RNA diferentes.

Quadro 14.5 Diferenças entre siRNAs e miRNAs.

Característica	siRNA	miRNA
Origem	mRNA, transpóson ou vírus	RNA transcrito a partir de gene diferente
Clivagem de	Duplex de RNA ou RNA de fita única que forma longos grampos de cabelo	RNA de fita única que forma grampos de cabelo curtos de RNA de fita dupla
Tamanho	21 a 25 nucleotídios	21 a 25 nucleotídios
Ação	Degradação do mRNA, inibição da transcrição, modificação da cromatina	Degradação do mRNA, inibição da tradução, modificação da cromatina
Alvo	Genes a partir dos quais eles foram transcritos	Genes diferentes dos quais eles foram transcritos

Em geral, os siRNAs têm complementaridade exata com seu mRNA-alvo ou sequências de DNA, enquanto os miRNAs têm complementaridade limitada com seus mRNAs-alvo. Os pequenos RNAs de interferência suprimem a expressão gênica ao degradar o mRNA ou inibir a transcrição, enquanto os miRNAs suprimem a expressão gênica ao inibir a tradução. Finalmente, os miRNAs em geral silenciam genes diferentes a partir dos quais os miRNAs foram transcritos, enquanto os siRNAs silenciam os genes a partir dos quais os siRNAs foram transcritos. Observe, entretanto, que essas diferenças entre siRNAs e miRNAs não são obrigatórias, e os cientistas estão encontrando pequenos RNAs que exibem as características de ambos. Por exemplo, alguns miRNAs (como os das plantas) têm complementaridade exata com sequências de mRNA e rompem essas sequências, características que estão associadas aos siRNAs.

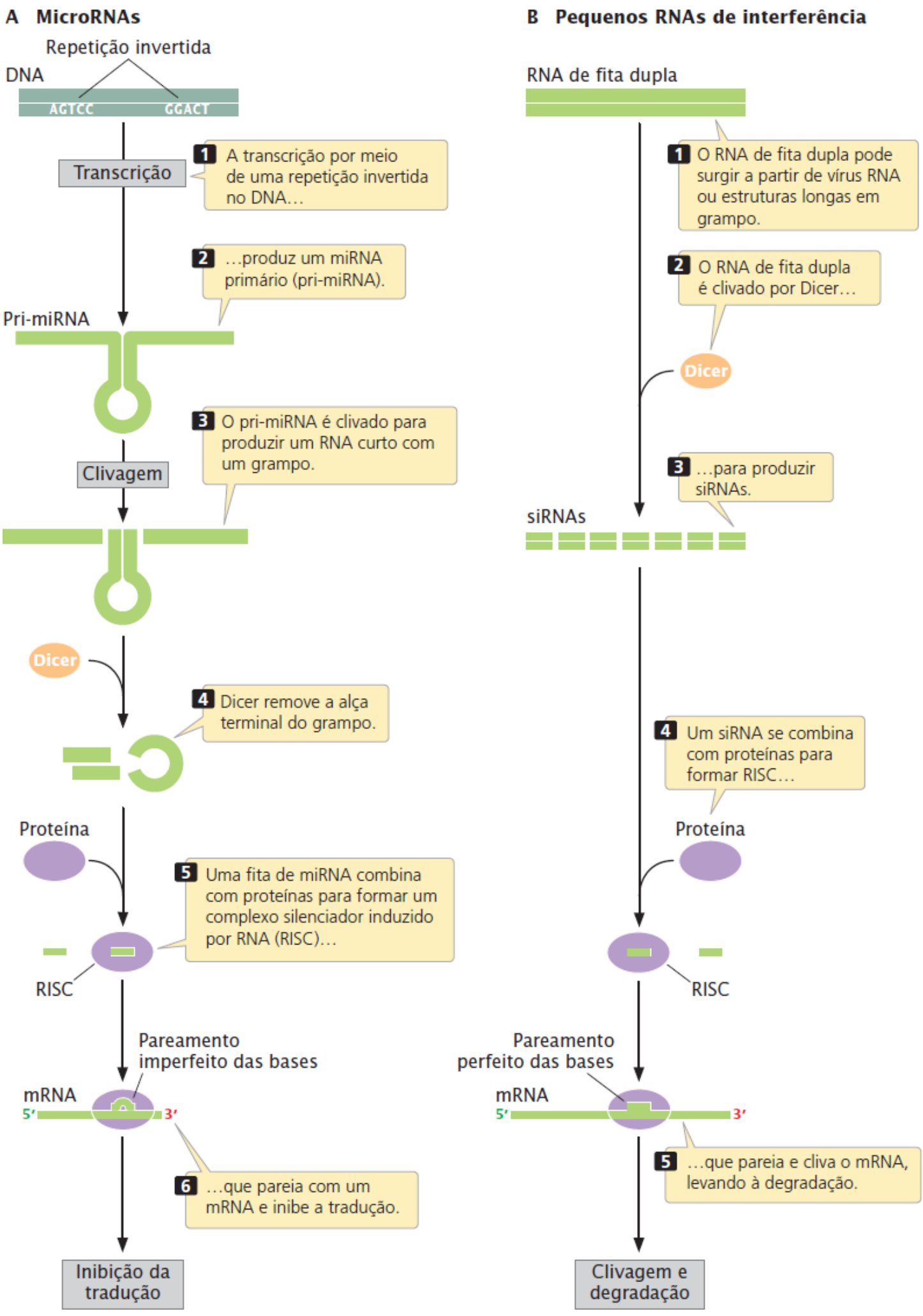

A MicroRNAs
Repetição invertida
DNA
AGTCC
GGACT
Transcrição
1 A transcrição por meio de uma repetição invertida no DNA…
2 …produz um miRNA primário (pri-miRNA).
Pri-miRNA
3 O pri-miRNA é clivado para produzir um RNA curto com um grampo.
Clivagem
Dicer
4 Dicer remove a alça terminal do grampo.
Proteína
5 Uma fita de miRNA combina com proteínas para formar um complexo silenciador induzido por RNA (RISC)…
RISC
Pareamento imperfeito das bases
mRNA
5′
3′
6 …que pareia com um mRNA e inibe a tradução.
Inibição da tradução
B Pequenos RNAs de interferência
RNA de fita dupla
1 O RNA de fita dupla pode surgir a partir de vírus RNA ou estruturas longas em grampo.
2 O RNA de fita dupla é clivado por Dicer…
Dicer
3 …para produzir siRNAs.
siRNAs
4 Um siRNA se combina com proteínas para formar RISC…
Proteína
RISC
Pareamento perfeito das bases
mRNA
5′
3′
5 …que pareia e cliva o mRNA, levando à degradação.
Clivagem e degradação

Figura 14.21 Pequenos RNAs interferentes e microRNAs são produzidos a partir de RNAs de fita dupla.

As moléculas de siRNA e miRNA combinam-se com as proteínas para formar um **complexo silenciador induzido por RNA** (RISC, ver Figura 14.21). A chave para o funcionamento do RISC é uma proteína chamada Argonauta. O RISC pareia com uma molécula de mRNA que tem uma sequência complementar com seu componente siRNA ou miRNA e cliva o mRNA, levando à sua degradação, ou reprime a tradução do mRNA. Alguns siRNAs também servem como guias para a metilação das sequências no DNA, e outros alteram a estrutura da cromatina; ambos afetam a transcrição.

Foram encontrados microRNAs em todos os organismos eucarióticos examinados até o momento, assim como os vírus: eles controlam a expressão dos genes que participam de muitos processos biológicos, incluindo crescimento, desenvolvimento e metabolismo. Os seres humanos têm mais de 450 miRNAs diferentes; os cientistas estimam que mais de um terço de todos os genes humanos seja regulado por miRNAs. A maioria dos genes de miRNAs é encontrada em regiões de DNA não codificado ou dentro de íntrons de outros genes.

Processamento e função de miRNAs e siRNAs. Os genes que codificam os miRNAs são transcritos em precursores maiores, chamados miRNA primário (pri-miRNA), que variam de muitas centenas a centenas de milhares de nucleotídios de comprimento (ver Figura 14.21 A). O pri-miRNA é clivado em uma ou mais moléculas de RNA menores com uma estrutura em grampo de cabelo. O Dicer se liga a essa estrutura de grampo de cabelo e remove a alça terminal. Uma das fitas de miRNA é incorporada ao RISC; a outra fita é liberada e degradada.

O RISC se fixa a uma sequência complementar no mRNA, em geral na região 3′ não traduzida do mRNA. A região de complementaridade próxima, chamada de região *seed*, é muito curta, em geral apenas sete nucleotídios de comprimento. Como a sequência *seed* é tão curta, cada miRNA pode potencialmente parear com sequências de centenas de diferentes miRNAs. Além disso, uma única molécula mRNA pode ter múltiplos sítios de ligação de miRNA. A inibição da tradução pode exigir a ligação de vários complexos RISC à mesma molécula mRNA.

Os pequenos RNAs de interferência são processados de modo semelhante (ver Figura 14.21 B). O RNA de fita dupla dos vírus ou longos grampos de cabelo são clivados por Dicer para produzir siRNAs, que se combinam com proteínas para formar um RISC. O RISC então pareia com as sequências no mRNA-alvo e cliva o mRNA, após o qual o mRNA é degradado.

Conceitos

Pequenos RNAs de interferência e microRNAs são RNAs minúsculos produzidos quando moléculas de RNA de fita dupla, maiores, são clivadas pela enzima Dicer. Os pequenos RNAs de interferência e os microRNAs participam de vários processos, incluindo a degradação de mRNA, a inibição da tradução, a metilação de DNA e a remodelagem da cromatina.

Checagem dos conceitos 9

De que modo siRNAs e miRNAs têm como alvo mRNAs específicos para degradação ou para repressão da tradução?

RNAs que interagem com Piwi

Os RNAs que interagem com Piwi (piRNAs) foram descobertos em 2001. Eles são mais longos que os siRNAs e os miRNAs, compostos por 24 a 30 nucleotídios, e são derivados de longos RNAs transcritos de fita única, ao contrário dos siRNAs e dos miRNAs, que são processados a partir de RNA de fita dupla. Também de maneira diferente dos siRNAs e dos miRNAs, Dicer não tem participação na produção de piRNAs.

Os RNAs que interagem com Piwi se combinam com proteínas Piwi que estão relacionadas com Argonauta, e suprimem a expressão e o movimento dos transpósons nas células germinativas dos animais. Embora o mecanismo de silenciar o transpóson por piRNAs ainda não esteja totalmente compreendido, sabemos que ele inclui a degradação do mRNA transcrito a partir dos transpósons, mudanças na estrutura da cromatina, que inibem a transcrição dos transpósons, e inibição da tradução de proteínas codificadas pelos transpósons. Não se sabe se os piRNAs têm alguma função fora das células germinativas.

CRISPR RNA

Após a descoberta de pequenos RNAs nos eucariotos, pequenos RNAs semelhantes chamados de CRISPR RNAs (crRNAs) foram descobertos nos procariotos: crRNAs são codificados por sequências de DNA encontradas nos genomas de bactérias e archaea chamadas de repetições palindrômicas curtas regularmente agrupadas (CRISPRs). As sequências palindrômicas leem o mesmo de trás para frente e de frente para trás nas duas fitas de DNA complementares. As CRISPRs são compostas por uma

série de tais sequências palindrômicas, separadas por sequências únicas que são homólogas para o DNA do bacteriófago ou dos genomas dos plasmídios. Por exemplo, na bactéria *Pseudomonas auruginosa*, as repetições palindrômicas de CRISPR têm 28 pb de comprimento, separadas por espaçadores de 32 pb.

CRISPR RNAs atuam na defesa contra a invasão de DNAs estranhos específicos, com os DNAs oriundos de bacteriófagos e plasmídios (ver *Capítulo 9*). Uma vez que elas têm como alvo moléculas específicas de DNA, o sistema CRISPR é comparado com o sistema imunológico dos vertebrados.

Por meio de um mecanismo ainda pouco compreendido, em que um bacteriófago ou plasmídio invade uma célula procariótica, pequenas porções do genoma invasor são inseridas como espaçadores entre as repetições palindrômicas em CRISPR (**Figura 14.22 A**). O DNA espaçador serve como memória desse DNA invasor. A região CRISPR é transcrita em um longo RNA precursor de CRISPR (**Figura 14.22 B**), que é então clivado pelas proteínas associadas a CRISPR (proteínas Cas) em crRNAS, cada um composto por uma sequência de espaçador (homóloga ao DNA invasor) flanqueada por uma parte da sequência palindrômica (**Figura 14.22 C**). Quando um DNA estranho oriundo da mesma fonte entra na célula, os crRNAs pareiam com ele e ajudam na clivagem desse DNA invasor (**Figura 14.22 D**). Desta maneira, os crRNAs servem como um sistema de defesa adaptado do RNA contra invasores.

Conceitos

RNAs que interagem com Piwi são encontrados nas células germinativas dos animais e inibem os transpósons. CRISPR RNAs são encontrados nos procariotos, onde atuam na defesa contra DNA estranho.

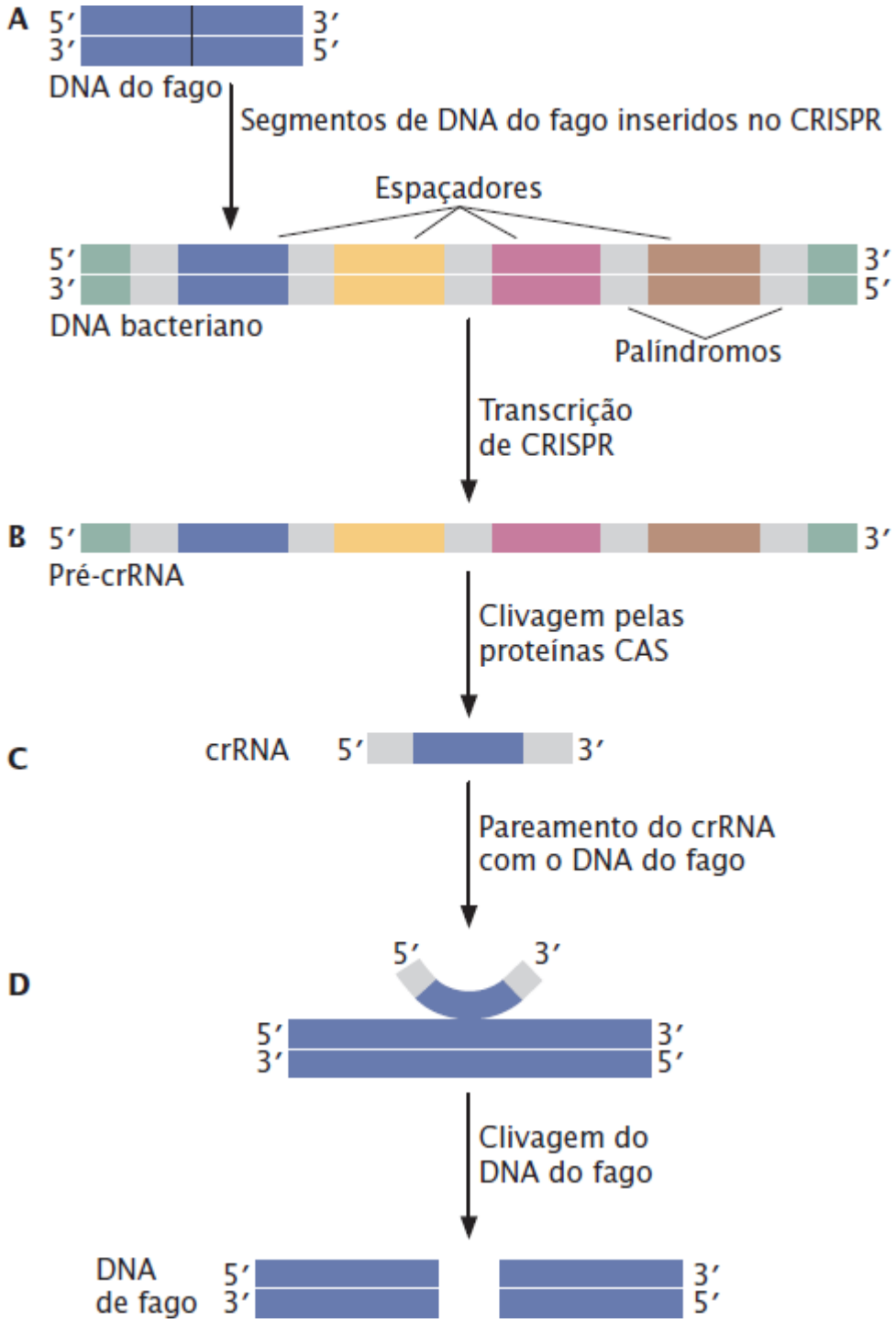


Figura 14.22 Os CRISPR RNAs atuam na defesa contra a invasão de DNA estranhos, como o DNA de bacteriófago e plasmídios.

14.6 Os RNAs longos não codificadores regulam a expressão gênica

Por muitos anos, nosso conhecimento sobre o RNA era limitado às moléculas com papel central na síntese das proteínas: mRNAs, tRNAs e rRNAs. Posteriormente, pequenos RNAs nucleares que participam no processamento pós-transcrição do RNA (snRNAs e snoRNAs) foram adicionados a essa lista. Começando no final dos anos 1990, os geneticistas passaram a identificar que vários RNAs pequenos (siRNAs, miRNAs, piRNAs e crRNAs) também eram abundantes e fundamentalmente importantes para o funcionamento da célula. Recentemente, tornou-se claro que a maioria dos genomas eucarióticos é transcrita – embora apenas 1% do genoma humano codifique diretamente as proteínas, mais de 80% são transcritos, produzindo muitas moléculas de RNA longas que não codificam as proteínas. Chamados de RNAs não codificadores longos (lncRNAs), esses RNAs têm mais de 100 nucleotídios de comprimento e não apresentam um *frame* de leitura (uma sequência com códons de início e de parada, que é traduzido pelos ribossomos). Foram descobertos milhares de lncRNAs nos últimos 5 anos. As sequências de DNA que os codificam, junto com outros DNAs de função desconhecida, são chamados de “matéria escura do genoma”.

Embora a função de muitos lncRNAs ainda não esteja clara, existe um indício crescente de que pelo menos eles exercem um

papel no controle da expressão gênica. Alguns lncRNAs interagem com proteínas que regulam a transcrição. Por exemplo, um lncRNA chamado lincRNA-p21 interage com uma proteína chamada p53, um fator de transcrição que ativa numerosos genes, incluindo os envolvidos no controle do ciclo celular e no câncer. Ao reprimir a p53, o lincRNA-p21 afeta a transcrição de centenas de genes. Outros lncRNAs modificam a estrutura da cromatina, que também regula a transcrição (ver *Capítulo 17*). Alguns lncRNAs têm sítios que são reconhecidos pelos miRNAs, e os lncRNAs servem como armadilhas para a fixação do miRNA. Assim, lncRNAs e mRNAs competem por um número limitado de miRNAs que regulam a tradução e a degradação um do outro. Outros lncRNAs ainda são complementares a sequências de mRNA e atuam no pareamento de base com o mRNA e evitam a tradução ou *splicing*.

Um dos lncRNAs mais bem estudados é o Xist RNA, que tem um papel central na compensação da dosagem nas células dos mamíferos (ver *Capítulo 4*). Para equilibrar a expressão dos genes ligados ao X nos machos (com um cromossomo X) e nas fêmeas (com dois cromossomos X), um dos cromossomos X em cada célula da fêmea dos mamíferos é inativado. É aleatório o cromossomo X que será inativado, e isso ocorre no início do desenvolvimento; uma vez desativado, esse cromossomo permanece inativo durante múltiplos ciclos da divisão celular. O Xist RNA é transcrito apenas a partir do cromossomo X destinado a ser inativo; o Xist RNA o reveste e recruta proteínas que metilam as histonas na cromatina. A metilação da cromatina, então, leva à inibição da transcrição dos genes no cromossomo X inativo. Pelo menos dois lncRNAs adicionais atuam na regulação da expressão do Xist RNA.

Evidências sugerem que outros lncRNAs também produzem o *imprinting* genômico (ver *Capítulos 4* e *21*). O *imprinting* ocorre quando um gene é expresso de modo diferente dependendo se ele é herdado do genitor ou da genitora. Muitos grupos de genes imprintados (*imprinted genes*) têm sequências codificadoras de lncRNAs, e as evidências sugerem que alguns genes imprintados sejam controlados por lncRNAS.

Conceitos

Os longos RNA não codificadores são longas moléculas de RNA que não codificam proteínas. As evidências sugerem que muitas dessas moléculas atuam no controle da expressão gênica.

Resumo dos conceitos

- Um gene é definido como uma sequência de nucleotídios de DNA transcritos em uma única molécula de RNA
- Os íntrons – sequências não codificadoras que interrompem as sequências codificadoras (éxons) dos genes – são comuns nas células eucarióticas, mas raras nas células bacterianas
- Uma molécula de mRNA tem três partes principais: uma região 5′ não traduzida, uma sequência codificadora de proteína e uma região 3′ não traduzida
- O mRNA bacteriano é traduzido imediatamente após a transcrição e sofre pouco processamento. O pré-mRNA do gene codificador de proteína eucariótico é muito processado: um nucleotídio modificado e um grupo metila, coletivamente chamados de cap, são adicionados à extremidade 5′ do pré-mRNA; a extremidade 3′ é clivada e uma cauda poli(A) é adicionada; e os íntrons são removidos. Os íntrons são removidos dentro de uma estrutura chamada spliceossomo, composta por diversos e pequenos RNAs nucleares e proteínas
- Alguns íntrons encontrados nos genes de mRNA e nos genes mitocondriais são *auto-splicing*
- Alguns pré-mRNAs sofrem *splicing* alternativo, no qual diferentes combinações de éxons são recompostas ou diferentes sítios de clivagem 3′ são usados
- Os RNAs mensageiros podem ser alterados por adição, deleção ou modificação dos nucleotídios na sequência codificadora, um processo chamado de edição do RNA
- Os RNAs transportadores (ou de transferência), que se prendem aos aminoácidos, são pequenas moléculas que assumem uma estrutura secundária comum e têm bases modificadas
- Ribossomos, os sítios de síntese de proteínas, são compostos por várias moléculas de RNA ribossômico e numerosas proteínas
- Pequenos RNAs de interferência, microRNAs, RNAs que interagem com Piwi e CRISPR RNAs têm participação importante no silenciamento gênico e em vários outros processos biológicos
- Os longos RNA não codificadores são moléculas de RNA que não codificam proteínas. Evidências sugerem cada vez mais que muitas dessas moléculas atuam no controle da expressão gênica.

Termos importantes

Anticódon
Base modificada

Cap 5′
Cauda poli(A)
Códon
Colinearidade
Complexo silenciador induzido por RNA (RISC)
Edição do RNA
Enzima modificadora de tRNA
Estrutura de folha de trevo
Éxon
Interferência por RNA (RNAi)
Íntron
Íntron de grupo I
Íntron de grupo II
Íntron de RNA transportador
Íntron do pré-mRNA nuclear
Lariat (laço)
Ponto de ramificação
Região 3′ não traduzida (3′ UTR)
Região 5′ não traduzida (5′ UTR)
Região codificadora de proteína
RNA guia
Sequência de Shine-Dalgarno
Sítios de clivagem 3′ múltiplos
Sítio de *splicing* 3′
Sítio de *splicing* 5′
Spliceossomo
Splicing ou recomposição alternativa
Splicing ou recomposição do RNA
Subunidade ribossômica maior
Subunidade ribossômica menor
Trans-splicing
Via de processamento alternativa

Respostas da Checagem dos conceitos

1. Quando o DNA foi hibridizado para o mRNA transcrito a partir deste, as regiões do DNA que não correspondiam ao RNA formavam alças.
2. Os íntrons do grupo I, do grupo II e íntrons pré-mRNA nucleares e íntrons de RNA transportador.
3. Uma proteína que adiciona o cap 5′ está associada à RNA polimerase II, que transcreve os pré-mRNAs, mas está ausente das RNA polimerases I e III, que transcrevem rRNA e tRNAs.
4. b.
5. c.
6. RNA guia.
7. b.
8. d.
9. Um siRNA ou miRNA combina com proteínas para formar RISC, que então se liga com mRNA por meio do pareamento complementar entre as bases no siRNA ou miRNA e as bases no mRNA.

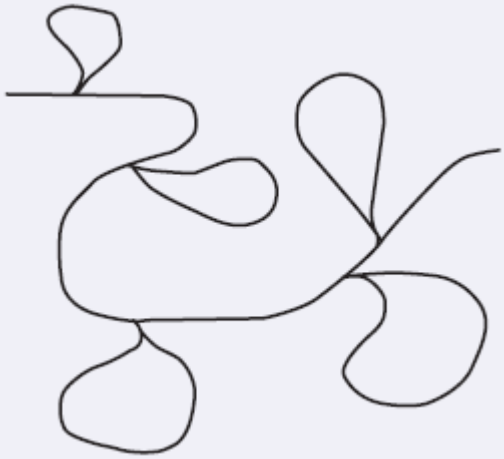
Problemas desenvolvidos

Problema 1

O DNA de um gene eucariótico foi isolado, desnaturado e hibridizado para o mRNA transcrito do gene; a estrutura hibridizada foi então observada com um microscópio eletrônico. Observou-se a estrutura ao lado.

a. Quantos íntrons e éxons há neste gene? Explique sua resposta.

b. Identifique os éxons e íntrons nesta estrutura hibridizada.



Estratégias para a solução

De quais informações você precisa para solucionar o problema?

- O número de íntrons e éxons e como você chegou à sua resposta
- A localização dos íntrons e éxons marcados na figura.

Quais informações são fornecidas para solucionar o problema?

- O DNA e o mRNA são oriundos de um eucarioto
- O DNA foi desnaturado e hibridizado para gerar o mRNA
- Uma imagem da estrutura hibridizada.

Para a solução deste problema, revise:

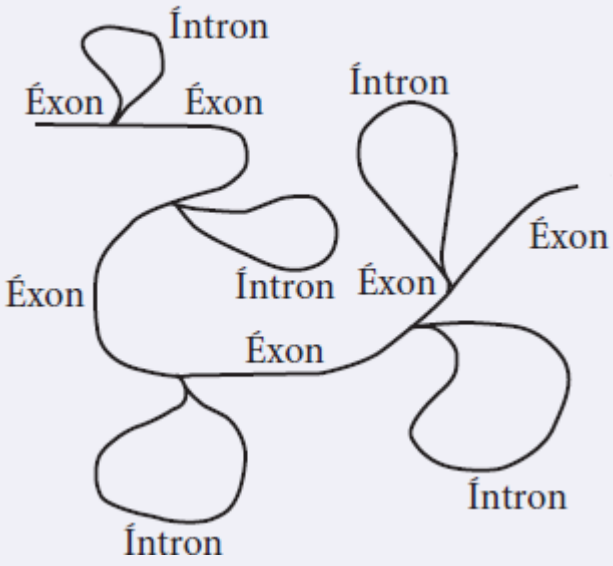
Íntrons, na Seção 14.1, e Figura 14.2 .

Etapas para a solução

a. Cada alça representa uma região na qual as sequências no DNA não têm sequências correspondentes no RNA; estas regiões são os íntrons. Existem cinco alças na estrutura hibridizada; então, existem obrigatoriamente cinco íntrons no DNA e seis éxons.

Lembrete: Íntrons são sequências não codificadoras encontradas nos genes eucarióticos.

b.



Dica: O número de íntrons será o número de éxons menos um.

Problema 2

Desenhe um mRNA bacteriano típico e o gene a partir do qual ele foi transcrito. Identifique as extremidades 5′ e 3′ das moléculas de RNA e DNA, assim como as regiões e sequências a seguir:

a. Promotor.
b. Região 5′ não traduzida.
c. Região 3′ não traduzida.
d. Sequência codificadora de proteína.
e. Sítio de início de transcrição.
f. Terminador.
g. Sequência de Shine-Dalgarno.
h. Códons de início e de parada.

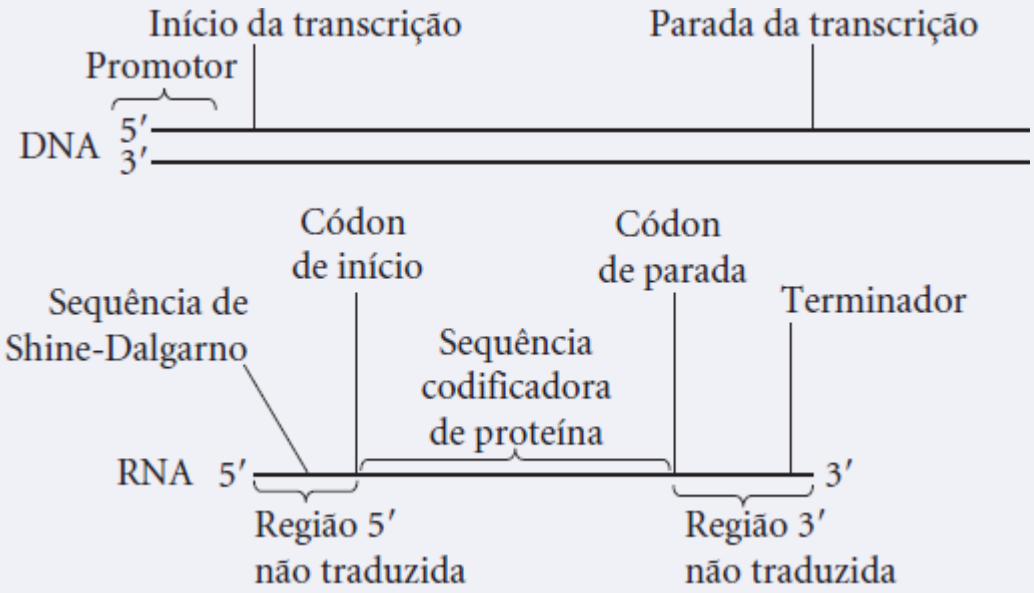
Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Um desenho do mRNA e do gene a partir do qual ele foi transcrito. As extremidades 5′ e 3′ das moléculas de mRNA e DNA. Localizações das estruturas listadas no desenho.

Quais informações são fornecidas para solucionar o problema?
- O gene é bacteriano
- Diferentes partes do DNA e do RNA que foram marcados.

Para a solução deste problema, revise:
Molde, na Seção 13.2, e *A estrutura do RNA mensageiro*, na Seção 14.2.

Etapas para a solução



Dica: Revise a estrutura de uma unidade de transcrição na Figura 13.6 e a estrutura do mRNA na Figura 14.5.

Questões de compreensão

Seção 14.1

1. O que é o conceito de colinearidade? De que modo esse conceito se aplica nas células bacterianas e eucarióticas?
2. Quais são as características dos íntrons?
3. Quais são os quatro tipos básicos de íntrons? Em quais organismos eles são encontrados?

Seção 14.2

4. Quais são os três principais elementos nas sequências de mRNA nas células bacterianas?
5. Qual é a função da sequência consenso de Shine-Dalgarno?
6. (a) O que é o cap 5′? (b) Como o cap 5′ é adicionado ao pré-mRNA eucariótico? (c) Qual é a função do cap 5′?
7. Como a cauda poli(A) é adicionada ao pré-mRNA? Qual é a finalidade da cauda poli(A)?
8. O que compõe o spliceossomo? Qual é a função do spliceossomo?
9. Explique o processo do *splicing* do pré-mRNA nos genes nucleares.
10. Descreva dois tipos de vias de processamento alternativas. Como elas levam à produção de múltiplas proteínas a partir de um único gene?
11. O que é edição do RNA? Explique o papel dos RNAs guias na edição do RNA.
12. Resuma os diferentes tipos de processamento que podem ocorrer no pré-mRNA.

Seção 14.3

13. Quais são as modificações no tRNA que ocorrem por meio do processamento?

Seção 14.4

14. Descreva a estrutura básica dos ribossomos nas células bacterianas e eucarióticas.
15. Explique como o rRNA é processado.

Seção 14.5

16. Qual é a origem dos pequenos RNAs de interferência, microRNAs e RNA que interagem com Piwi? Qual é a função dessas moléculas de RNA na célula?
17. Quais são as semelhanças e as diferenças entre siRNAs e miRNAs?
18. Como os miRNAs são processados?

Questões e problemas aplicados

Seção 14.1

*19. A distrofia muscular de Duchenne é causada por uma mutação em um gene com 2,5 milhões de nucleotídios e especifica uma proteína chamada distrofina. Entretanto, menos de 1% do gene codifica de fato os aminoácidos na proteína distrofina. Com base no que você agora sabe sobre a estrutura do gene e o processamento do RNA nas células eucarióticas, forneça uma possível explicação para o tamanho grande do gene distrofina.
20. O que aconteceria no experimento ilustrado na **Figura 14.2** se o DNA e o RNA que foram misturados viessem de diferentes organismos, por exemplo, de um verme e de um porco?
21. Para o gene da ovalbumina apresentado na **Figura 14.3**, onde estariam as regiões 5′ não traduzida e 3′ não traduzida no DNA e no RNA?

Seção 14.2

22. Quais são as diferenças entre os mRNAs das células bacterianas e dos pré-mRNAs das células eucarióticas? Quais são as diferenças entre os mRNAs maduros das células bacterianas e eucarióticas?
23. As regiões 5′ não traduzidas (5′ UTR) dos mRNAs eucarióticos são codificadas por sequências no promotor, éxon ou íntron do gene? Explique sua resposta.
*24. Desenhe um típico gene eucariótico e o pré-mRNA e mRNA derivados deste. Considere que o gene tem três éxons. Identifique os itens a seguir e, para cada item, forneça uma breve descrição de sua função:
    a. Região 5′ não traduzida.
    b. Promotor.
    c. Sequência consenso AAUAAA.
    d. Sítio de início de transcrição.
    e. Região 3′ não traduzida.
    f. Íntrons.
    g. Éxons.
    h. Cauda poli(A).
    i. Cap 5′.

25. Como a deleção da sequência de Shine-Dalgarno afetaria um mRNA bacteriano?
26. Qual seria o efeito mais provável de deslocar a sequência consenso AAUAAA apresentada na Figura 14.7 dez nucleotídios *upstream*?
27. Como a deleção das sequências ou características a seguir mais provável afetaria um pré-mRNA eucariótico?
    a. Sequência consenso AAUAAA.
    b. Cap 5′.
    c. Cauda poli(A).
28. Suponha que haja uma mutação no meio de um grande íntron de um gene codificador de uma proteína. Qual seria o efeito mais provável da mutação na sequência de aminoácidos desta proteína? Explique sua resposta.

*29. Um geneticista induz uma mutação em uma linhagem de células cultivada em laboratório. A mutação ocorre em um dos genes que codifica proteínas participantes da clivagem e da poliadenilação do mRNA eucariótico. Qual seria o efeito imediato desta mutação nas moléculas de RNA nas células cultivadas?

30. Um geneticista faz uma mutação no gene para proteínas que se ligam à cauda poli(A) em uma linhagem de células cultivada em laboratório. Qual seria o efeito imediato desta mutação nas células cultivadas?
31. Um geneticista isola um gene com oito éxons. Ele isola o mRNA maduro produzido por este gene. Após isolar o DNA de fita única, ele mistura o DNA de fita única com o RNA. Parte do DNA de fita única hibridiza (pareia) com o mRNA complementar. Desenhe como seriam os híbridos DNA-RNA à microscopia eletrônica.
32. Um geneticista descobre que duas proteínas diferentes são codificadas pelo mesmo gene. Uma proteína tem 56 aminoácidos, e a outra tem 82 aminoácidos. Indique uma possível explicação sobre como o mesmo gene consegue codificar ambas as proteínas.
33. Que conclusão você pode tirar sobre os tamanhos relativos das duas proteínas produzidas por *splicing* alternativo na Figura 14.12?

*34. Explique como cada um dos processos a seguir complica o conceito de colinearidade.
    a. *Trans-splicing*.
    b. *Splicing* alternativo.
    c. Edição do RNA.

Seção 14.5

35. A interferência do RNA pode ser deflagrada quando repetições invertidas são transcritas para uma molécula de RNA que, então, se dobra para formar o RNA de fita dupla. Escreva uma sequência de repetições invertidas dentro de uma molécula de RNA. Usando um diagrama, mostre como o RNA com as repetições invertidas pode se dobrar para formar um RNA de fita dupla.

36. No início dos anos 1990, Carolyn Napoli e seus colaboradores estavam trabalhando com petúnias, tentando criar, por meio de engenharia genética, uma variedade de pétalas púrpura-escuro pela introdução de numerosas cópias de um gene codificador de pétalas púrpura (C. Napoli, C. Lemieux e R. Jorgensen. 1990. *Plant Cell* 2:279-289). A base racional era que as cópias extras do gene levariam a produzir mais pigmento cor púrpura e resultariam em uma petúnia com um tom mais escuro de púrpura. Entretanto, para sua surpresa, muitas das plantas com cópias extras do gene da cor púrpura eram completamente brancas ou tinham apenas manchas coloridas. A análise molecular revelou que o nível de mRNA produzido pelo gene púrpura foi reduzido em 50 vezes nas plantas modificadas se comparado com os níveis de mRNA nas plantas do tipo selvagem. De alguma maneira, a introdução de cópias extras do gene púrpura silenciou as cópias introduzidas e os próprios genes da cor púrpura da planta. Dê uma possível explicação para como a introdução de numerosas cópias do gene púrpura silenciou todas as cópias deste gene.

Questões desafiadoras

Seção 14.2

37. O *splicing* alternativo ocorre em mais de 90% dos genes humanos que codificam proteínas. Os pesquisadores descobriram como a recomposição do pré-mRNA é afetada pela sequência promotora do pré-mRNA (D. Auboeuf *et al*. 2002. *Science* 298:416-419). Além disso, os fatores que afetam a taxa de prolongamento da RNA polimerase durante a transcrição afetam o tipo de *splicing* que ocorre. Esses achados sugerem que o processo de transcrição afeta o *splicing*. Proponha um

ou mais mecanismos que explicariam como a transcrição poderia afetar o *splicing* alternativo.

38. A distrofia muscular de Duchenne (DMD) é uma doença genética recessiva ligada ao cromossomo X causada por mutações no gene que codifica a distrofina, uma grande proteína importante no desenvolvimento das fibras musculares normais. O gene da distrofina é imenso, incluindo 2,5 milhões de pares de base e 79 éxons e 78 íntrons. Muitas das mutações que provocam DMD produzem códons de paradas prematuras, que fazem com que a síntese de proteínas pare, gerando uma forma muito reduzida e não funcional de distrofina. Alguns geneticistas propuseram tratar os pacientes com DMD ao introduzir pequenas moléculas de RNA que fazem o spliceossomo pular o éxon que contém o códon de parada. A introdução de pequenos RNAs produzirá uma proteína que será um pouco encurtada (porque um éxon foi "pulado" e não existem alguns aminoácidos), mas ainda pode resultar em uma proteína com alguma função (A. Goyenvalle *et al*. 2004. *Science* 306:1796-1799). Os RNAs pequenos usados para "pular" o éxon são complementares às bases no pré-mRNA. Se você estivesse planejando pequenos RNAs para promover a omissão de éxons para o tratamento da DMD, quais sequências os pequenos RNAs teriam?
39. Nas células eucarióticas, uma cauda poli(A) é normalmente adicionada a moléculas de pré-mRNA, mas não a rRNA ou tRNA. Com o uso de técnicas de DNA recombinante, um gene codificador de proteína (que é normalmente transcrito pela RNA polimerase II) pode ser conectado a um promotor para RNA polimerase I. Esse gene híbrido é transcrito pela RNA polimerase I e o pré-mRNA adequado é produzido, mas esse pré-mRNA não é rompido na extremidade 3′ e uma cauda poli(A) não é adicionada.

    Proponha um mecanismo para explicar como o tipo de promotor encontrado na extremidade 5′ de um gene pode afetar se uma cauda poli(A) é adicionada à extremidade 3′.
40. As proteínas SR são essenciais para a montagem correta do spliceossomo e são conhecidas por participarem da regulação do *splicing* alternativo. Surpreendentemente, o papel das proteínas SR na seleção do sítio de *splicing* e do *splicing* alternativo é afetado pelo promotor usado para a transcrição do pré-mRNA. Por exemplo, por meio da engenharia genética podem ser criados promotores da RNA polimerase II com sequências um pouco diferentes. Quando pré-mRNAs com exatamente as mesmas sequências são transcritos a partir de dois diferentes promotores de RNA polimerase II com pouca diferença na sequência, qual promotor é usado pode afetar o modo como o pré-mRNA é recomposto.

    Proponha um mecanismo que explique como a sequência de DNA de um promotor de RNA polimerase II poderia afetar o *splicing* alternativo do pré-mRNA.

15 Código Genético e Tradução

Hutterites, ribossomos e a síndrome de Bowen-Conradi

A natureza essencial do ribossomo – a fábrica de proteínas da célula – é ilustrada de maneira simples pelas crianças com a síndrome de Bowen-Conradi. Com nariz proeminente, cabeça pequena e uma curvatura incomum do quinto dedo da mão, essas crianças não se desenvolvem nem ganham peso, e em geral morrem no primeiro ano de vida.

Quase todas as crianças com síndrome de Bowen-Conradi são hutterites, um ramo dos anabatistas, que surgiram nos anos 1500 nos Alpes do Tirol, na Áustria. Após anos de perseguição, os hutterites migraram para a Dakota do Sul nos anos de 1870 e se espalharam nos estados de pradaria vizinhos e nas províncias do Canadá. Atualmente, existem cerca de 40.000 hutterites na América do Norte. Eles vivem em fazendas coletivas, são muito pacifistas e raramente casam com pessoas de fora da comunidade.

A síndrome de Bowen-Conradi é herdada como um distúrbio autossômico recessivo, e a associação dessa síndrome com a comunidade hutterite é consequente a história genética única do grupo. O *pool* de genes dos hutterites de hoje na América do Norte pode ser traçado até menos de 100 pessoas que migraram para a Dakota do Sul no final do século 19. A incidência maior dessa síndrome nos hutterites atualmente é oriunda do efeito do fundador – a existência de um alelo que provoca a síndrome em um ou mais fundadores originais – e se espalhou à medida que os hutterites casaram entre si na comunidade. Em razão do efeito fundador e da consanguinidade (ver *Capítulo 25*), muitos hutterites atualmente são parentes próximos, como primos de primeiro grau. Essa relação genética próxima entre eles aumenta a probabilidade de que uma criança herde duas cópias do alelo recessivo e tenha a síndrome de Bowen-Conradi; de fato, quase 1 em 10 hutterites é portador heterozigoto do alelo que causa a doença.

Embora essa síndrome tenha sido descrita pela primeira vez em 1976, sua base genética e bioquímica era um mistério. Após uma busca de sete anos para encontrar o gene responsável, os pesquisadores da University of Manitoba determinaram em 2009 que a síndrome resulta da mutação de um único par de bases no gene *EMG1*, localizado no cromossomo 12.

A descoberta do gene para a síndrome de Bowen-Conradi deu uma percepção imediata da natureza bioquímica da doença. Embora pouco se saiba sobre a função do gene *EMG1* nos humanos, estudos iniciais em levedura revelaram que ele codifica uma proteína que auxilia na montagem do ribossomo. Como foi discutido no Capítulo 14, o ribossomo é composto por uma subunidade maior e uma menor. A subunidade menor nos humanos é composta por um único pedaço de RNA ribossômico conhecido como 18S rRNA, e várias proteínas. A proteína codificada pelo gene *EMG1* tem papel essencial no processamento do 18S rRNA e ajuda a montá-lo na subunidade menor do ribossomo. Por causa de uma mutação no gene *EMG1*, os recém-nascidos com a síndrome de Bowen-Conradi produzem ribossomos que funcionam mal, e o processo pelo qual todas as proteínas são sintetizadas é afetado.

A síndrome de Bowen-Conradi ilustra a extrema importância da tradução, o processo de síntese de proteínas, que é o foco deste capítulo. Começaremos examinando a relação molecular entre o genótipo e o fenótipo. A seguir, estudaremos o código genético – as instruções que especificam a sequência de aminoácidos de uma proteína – e então examinaremos o mecanismo de síntese de proteínas. Nosso primeiro ponto será a síntese de proteínas nas células bacterianas, mas vamos examinar algumas das diferenças nas células eucarióticas. Observaremos alguns aspectos adicionais sobre a síntese de proteínas no final do capítulo.

15.1 Muitos genes codificam proteínas

A primeira pessoa a sugerir a existência de uma relação entre genótipo e proteínas foi o médico inglês Archibald Garrod. Em 1908, Garrod propôs, de forma correta, que os genes codificam as enzimas, mas, infelizmente, sua teoria não impressionou muito seus contemporâneos. Não até os anos de 1940, quando George Beadle e Edward Tatum examinaram a base genética das vias bioquímicas na *Neurospora*, e a relação entre genes e proteínas se tornou amplamente aceita. O trabalho de Beadle e Tatum

ajudou a definir a relação entre genótipo e fenótipo ao criar a hipótese de um gene, uma enzima; a ideia que cada gene codifica uma enzima separada.

Hipótese de um gene, uma enzima

Beadle e Tatum usaram o bolor de pão *Neurospora* para estudar os resultados bioquímicos das mutações. A *Neurospora* é fácil de ser cultivada em laboratório, e a principal parte vegetativa do fungo é haploide, o que facilita observar os efeitos das mutações recessivas.

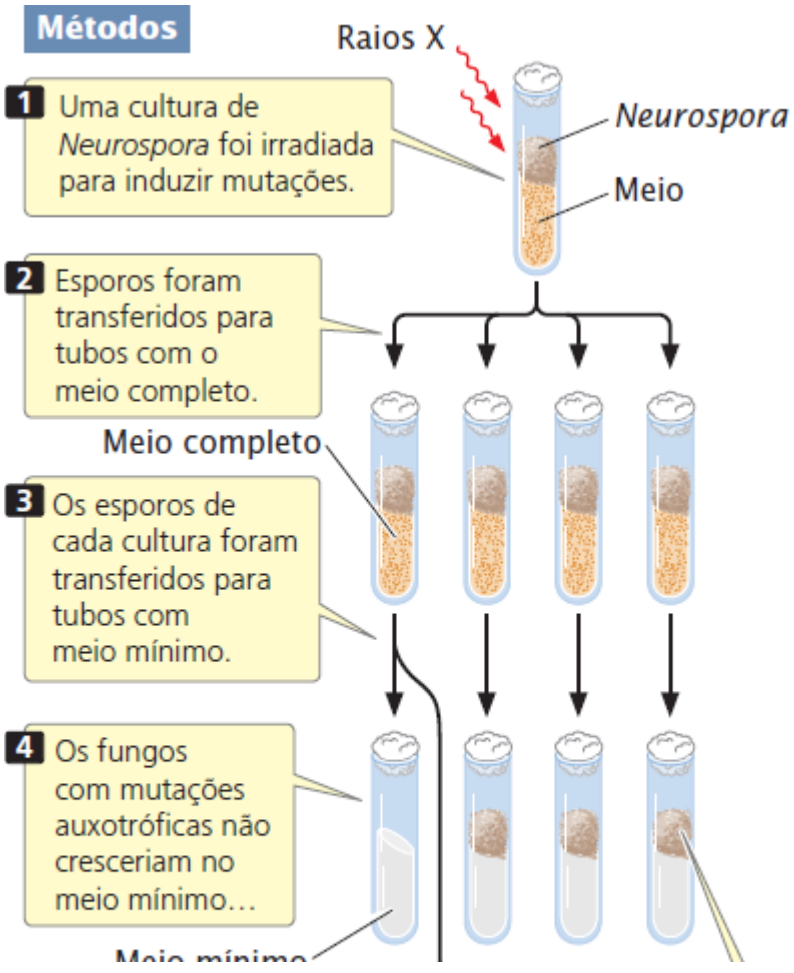
O tipo selvagem de *Neurospora* cresce no meio mínimo, que tem sais inorgânicos, nitrogênio, uma fonte de carbono como a sacarose e a vitamina biotina. O fungo pode sintetizar todas as moléculas biológicas que precisa a partir desses compostos básicos. Entretanto, podem surgir mutações que interrompam o crescimento do fungo ao destruir sua capacidade de sintetizar uma ou mais moléculas biológicas essenciais. Esses mutantes com deficiências nutricionais, chamados de auxótrofos (ver *Capítulo 9*), não crescem no meio mínimo, mas podem crescer em um meio com a substância que eles não conseguem mais sintetizar.

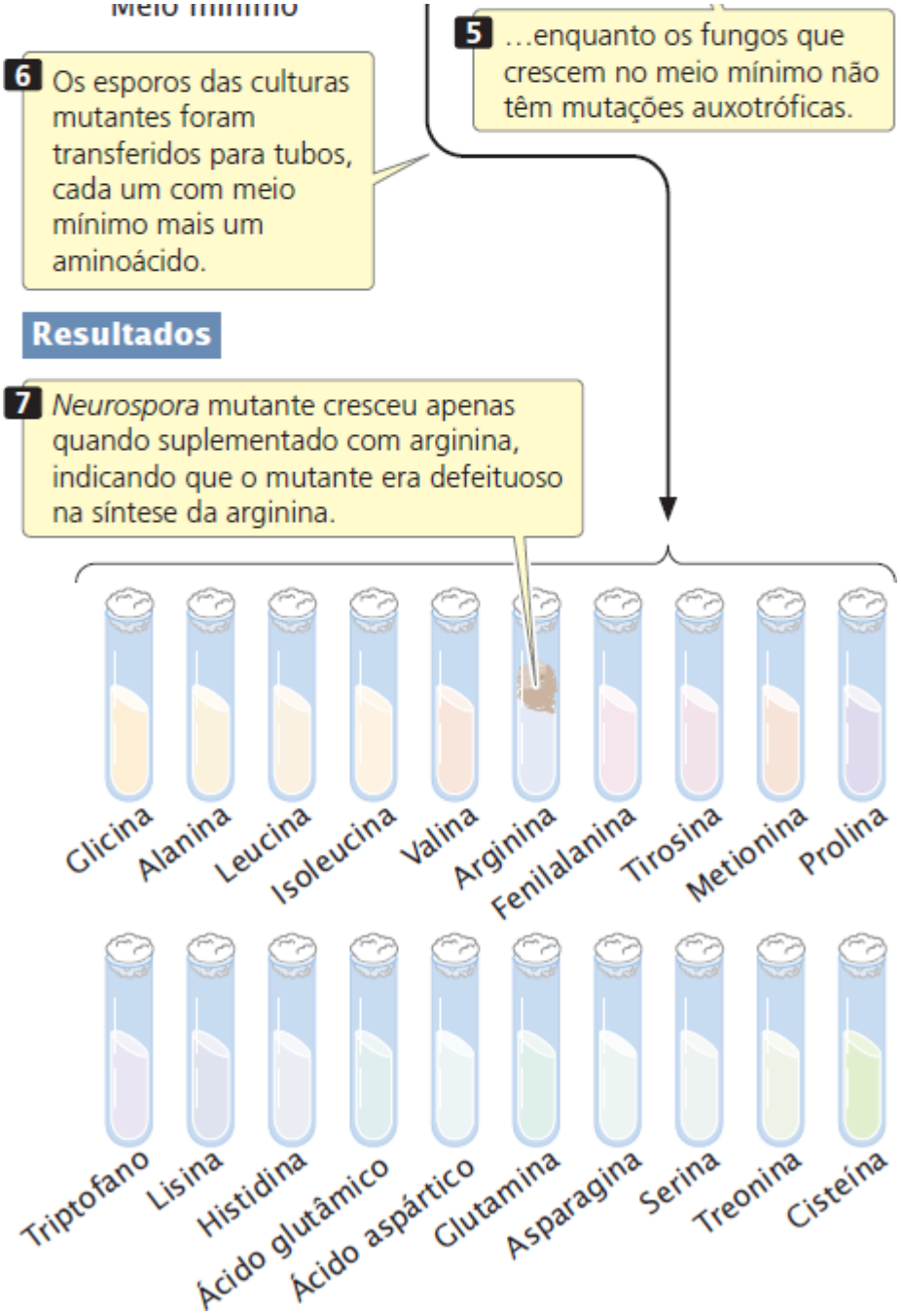
Beadle e Tatum primeiro irradiaram os esporos de *Neurospora* para induzir mutações (**Figura 15.1**). Então colocaram os esporos em diferentes tubos de cultura contendo meio completo (meio com todas as substâncias biológicas que precisam para crescer). Esses esporos viraram fungos e produziram esporos por mitose. A seguir, eles transferiram os esporos de cada cultura para tubos contendo o meio mínimo. Os fungos com mutações auxotróficas cresciam no meio completo, mas não cresciam no meio mínimo, o que possibilitou que Beadle e Tatum identificassem as culturas que tinham as mutações.

Após terem determinado que certa cultura tinha uma mutação auxotrófica, Beadle e Tatum começaram a determinar o *efeito* específico da mutação. Eles transferiram os esporos de cada cepa mutante do meio completo para vários tubos (ver Figura 15.1), cada um com o meio mínimo mais uma molécula de uma variedade de moléculas biológicas essenciais, como um aminoácido. Se os esporos cresciam em um tubo, Beadle e Tatum eram capazes de identificar a substância adicionada como a molécula biológica cuja sua síntese foi afetada pela mutação. Por exemplo, um mutante auxotrófico que cresceria apenas no meio mínimo ao qual foi adicionada arginina deve ter uma mutação que interrompe a síntese da arginina.

Experimento

Pergunta: Como as mutações auxotróficas podem ser isoladas e identificadas?





Conclusão: A mutação poderia ser identificada pela capacidade dos esporos em crescer em um meio mínimo suplementado com a substância que os esporos não conseguiam sintetizar.

Figura 15.1 Beadle e Tatum desenvolveram um método para isolar mutantes auxotróficos de *Neurospora*.

Adrian Srb e Norman H. Horowitz aplicaram pacientemente esse procedimento para dissecar geneticamente uma via bioquímica de múltiplas etapas da síntese da arginina (**Figura 15.2**). Primeiro, eles isolaram vários mutantes auxotróficos cujo crescimento exigia arginina. Então testaram a capacidade desses mutantes em crescer no meio mínimo suplementado com três compostos: ornitina, citrulina e arginina. A partir dos resultados, eles foram capazes de colocar os mutantes nesses três grupos com base em que as substâncias possibilitaram o crescimento (**Quadro 15.1**).

Com base nesses resultados, Srb e Horowitz propuseram que a via bioquímica que produzia o aminoácido arginina tinha pelo menos três etapas:

$$\text{Precursor} \xrightarrow{\text{Etapa 1}} \text{Ornitina} \xrightarrow{\text{Etapa 2}} \text{Citrulina} \xrightarrow{\text{Etapa 3}} \text{Arginina}$$

Eles concluíram que as mutações no grupo I afetam a etapa 1 desta via, as mutações no grupo II afetam a etapa 2 e as

mutações no grupo III afetam a etapa 3. Mas como eles sabiam que a ordem dos compostos na via bioquímica estava correta?

Observe que se a etapa 1 está bloqueada por uma mutação, então a adição de ornitina ou citrulina possibilita o crescimento, porque estes compostos ainda podem ser convertidos em arginina (ver Figura 15.2). Da mesma maneira, se a etapa 2 for bloqueada, a adição de citrulina possibilita o crescimento, mas a adição de ornitina não tem efeito. Se a etapa 3 for bloqueada, os esporos crescem apenas se for adicionada arginina ao meio. O princípio fundamental é que um mutante auxotrófico não pode sintetizar qualquer composto que surja após a etapa bloqueada por uma mutação.

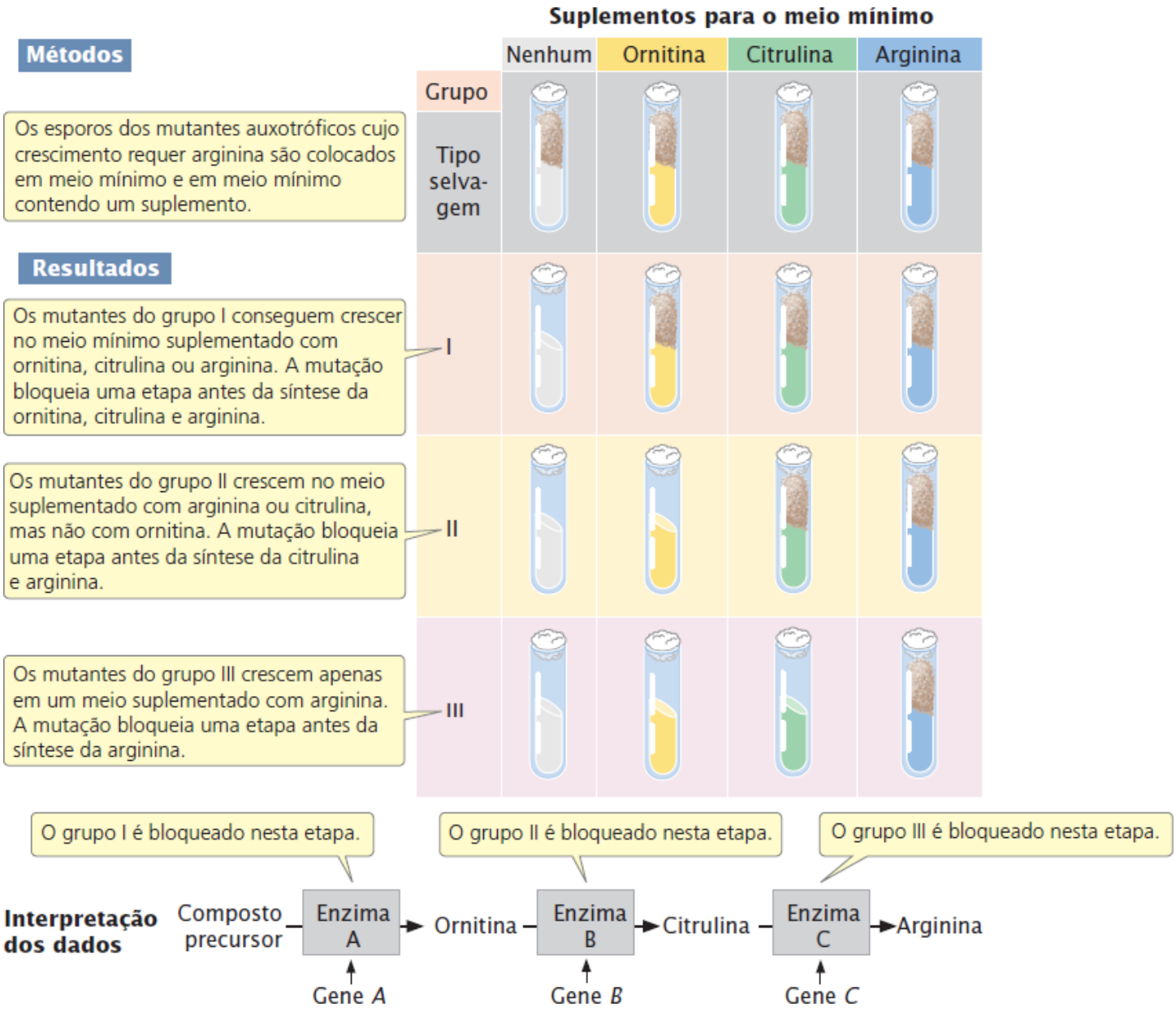
Quadro 15.1 Crescimento de mutantes auxotróficos da arginina no meio mínimo com vários suplementos.

Número da cepa mutante	Ornitina	Citrulina	Arginina
Grupo I	+	+	+
Grupo II	−	+	+
Grupo III	−	−	+

Nota: Um sinal mais (+) indica crescimento; um sinal menos (−) indica ausência de crescimento.

Experimento

Pergunta: O que os efeitos da mutação genética em uma via bioquímica nos dizem sobre a relação proteína — gene?



Conclusão: Cada gene codifica uma proteína separada – neste caso, uma enzima.

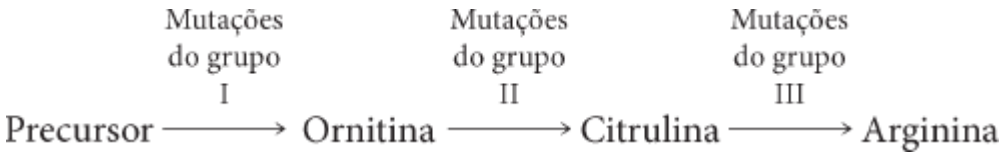
Figura 15.2 Método usado para determinar a relação entre os genes e enzimas na *Neurospora*. Essa via bioquímica leva à síntese de arginina em *Neurospora*. As etapas na via são catalisadas, por enzimas afetadas por mutações.

Usando esse raciocínio com as informações no Quadro 15.1, podemos ver que a adição de arginina ao meio torna possível que os três grupos de mutantes cresçam. Portanto, as etapas bioquímicas afetadas por todos os mutantes vêm antes da etapa que produz a arginina. A adição de citrulina possibilita que os mutantes dos grupos I e II cresçam, mas não os mutantes do grupo III; portanto, as mutações do grupo III devem afetar uma etapa bioquímica que ocorre após a produção da citrulina, mas antes da produção da arginina.

$$\text{Citrulina} \xrightarrow{\text{Mutações do grupo III}} \text{Arginina}$$

A adição de ornitina torna possível o crescimento dos mutantes do grupo I, mas não os mutantes dos grupos II e III; assim, as mutações nos grupos II e III afetam as etapas que vêm após a produção da ornitina. Já estabelecemos que as mutações do grupo II afetam uma etapa antes da produção da citrulina; então, as mutações do grupo II obrigatoriamente bloqueiam a conversão de ornitina em citrulina.

Como as mutações do grupo I afetam alguma etapa antes da produção da ornitina, podemos concluir que elas devem afetar a conversão de algum precursor em ornitina. Agora podemos esboçar a via bioquímica que produz ornitina, citrulina e arginina.



É importante destacar que este procedimento não detecta necessariamente todas as etapas em uma via; ele apenas detecta as etapas que produzem os compostos testados.

Usando as mutações e esse tipo de raciocínio, Beadle, Tatum e outros pesquisadores conseguiram identificar os genes que controlam várias vias biossintéticas em *Neurospora*. Eles estabeleceram que cada etapa em uma via é controlada por uma enzima diferente, como demonstrado na Figura 15.2 para a via da arginina. Eles também fizeram cruzamentos genéticos e mapearam os experimentos (ver *Capítulo 7*), e conseguiram demonstrar que as mutações que afetam qualquer uma das etapas de uma via sempre ocorriam na mesma localização no cromossomo. Beadle e Tatum justificaram que as mutações que afetam uma etapa bioquímica particular ocorriam em um único *locus* que codificava uma enzima particular. Essa ideia ficou conhecida como a **hipótese de um gene, uma enzima**: os genes codificam as enzimas, e cada gene codifica uma enzima separada. Embora os genes que Beadle e Tatum examinaram codificassem enzimas, muitos codificam proteínas que não são enzimas; então, em uma visão mais ampla, sua ideia era que cada gene codifica uma proteína. Quando as descobertas das pesquisas mostraram que algumas proteínas são compostas por mais de uma cadeia de polipeptídios e que diferentes cadeias de polipeptídios são codificadas por genes separados, esse modelo foi modificado para ser a **hipótese de um gene, um polipeptídio**. Resolva o Problema 16

Conceitos

Os estudos das vias bioquímicas de Beadle e Tatum no fungo *Neurospora* ajudaram a definir a relação entre genótipo e fenótipo ao levar para a hipótese de um gene, uma enzima, a ideia de que cada gene codifica uma enzima separada. Esta hipótese foi modificada posteriormente para a hipótese de um gene, um polipeptídio.

Checagem dos conceitos 1

A mutação auxotrófica 103 cresce em meio mínimo suplementado com A, B ou C; a mutação 106 cresce em meio suplementado com A e C, mas sem B, e a mutação 102 cresce apenas em meio suplementado com C. Qual é a ordem de A, B ou C na via bioquímica?

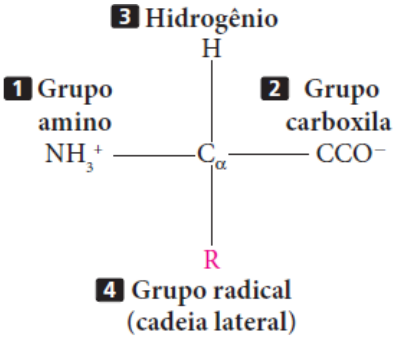
Estrutura e função das proteínas

As proteínas são peças cruciais para todos os processos nos seres vivos. Muitas proteínas são enzimas, os catalisadores biológicos que impulsionam as reações químicas da célula; outras são componentes estruturais, fornecendo o arcabouço e o suporte para membranas, filamentos, ossos e pelos. Algumas proteínas ajudam a transportar substâncias; outras têm função reguladora, de comunicação ou defesa.

Aminoácidos. Todas as proteínas são compostas por **aminoácidos**, ligados de uma ponta a outra. Existem 20 aminoácidos comuns nas proteínas; eles estão apresentados na **Figura 15.3** com suas abreviações de três letras ou uma letra (outros aminoácidos encontrados nas proteínas são formas modificadas dos aminoácidos comuns). Os 20 aminoácidos comuns têm estrutura semelhante: cada um tem um carbono central ligado a um grupo amino, um átomo de hidrogênio, um grupo carboxila e um grupo R (radical) diferente para cada aminoácido. Os aminoácidos nas proteínas são unidos por **ligações peptídicas** (**Figura 15.4**) para formar cadeias de **polipeptídios**, e uma proteína tem uma ou mais cadeias de polipeptídios. Como os ácidos nucleicos, os polipeptídios têm polaridade, com uma extremidade com um grupo amino livre (NH_3^+) e a outra com um grupo carboxila livre (COO^-). Algumas proteínas têm apenas alguns aminoácidos, enquanto outras podem ter milhares.

Estrutura da proteína. Como os ácidos nucleicos, a estrutura molecular das proteínas tem vários níveis de organização. A

estrutura primária de uma proteína é sua sequência de aminoácidos (**Figura 15.5 A**). Por meio das interações entre os aminoácidos vizinhos, uma cadeia de polipeptídios se dobra e gira em uma *estrutura secundária* (**Figura 15.5 B**); as duas estruturas secundárias comuns encontradas nas proteínas são a folha dobrada beta (*β)* e a hélice alfa (*α*). As estruturas secundárias interagem e se dobram mais para formar uma *estrutura terciária* (**Figura 15.5 C**), que é o formato geral, tridimensional da proteína. As estruturas secundárias e terciárias de uma proteína são determinadas em grande parte pela estrutura primária – a sequência de aminoácidos – da proteína. Finalmente, algumas proteínas têm duas ou mais cadeias de polipeptídios que se associam para produzir uma *estrutura quaternária* (**Figura 15.5 D**).



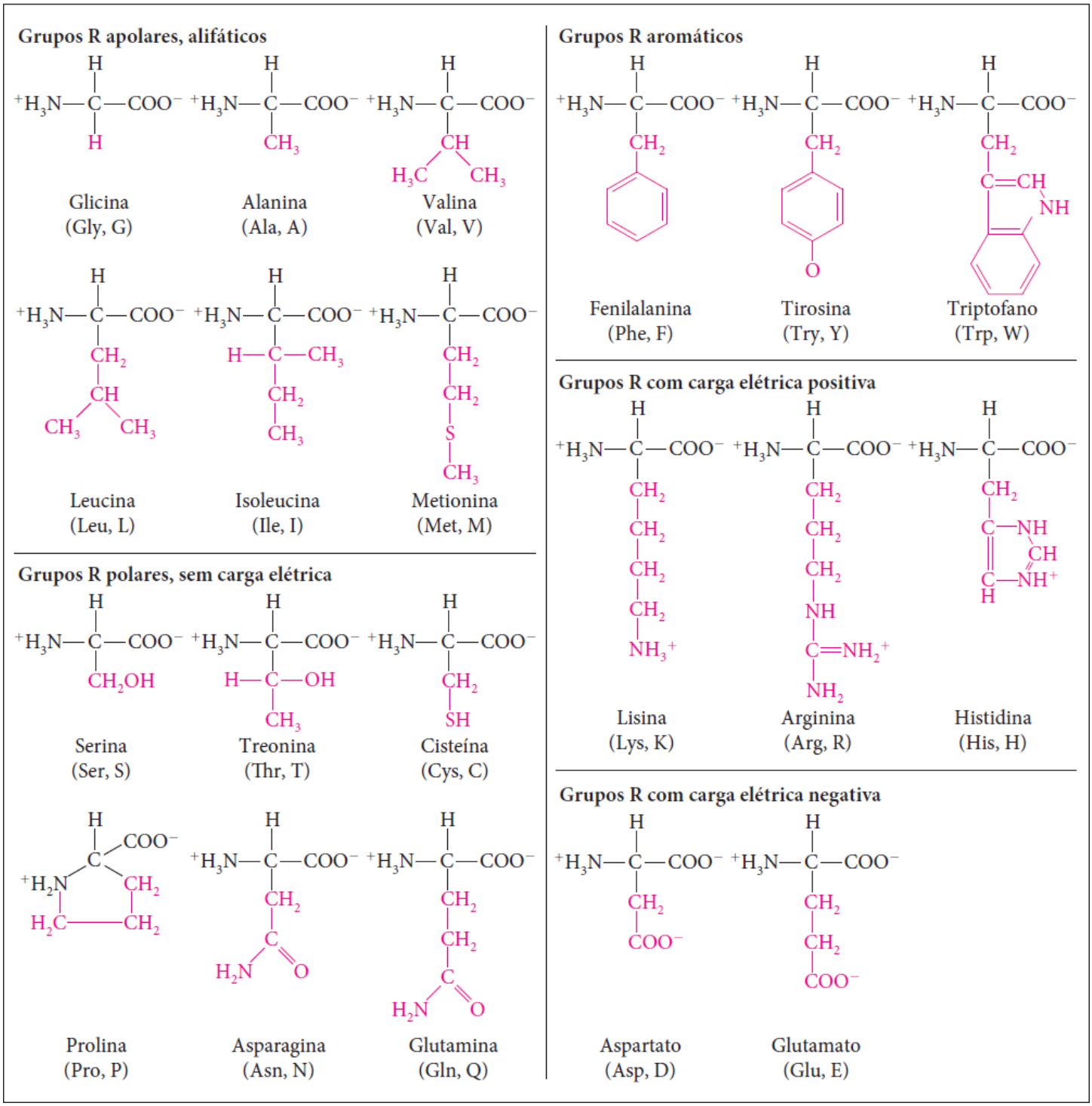


Figura 15.3 Os aminoácidos comuns têm estruturas semelhantes. Cada aminoácido tem um átomo de carbono central (C_α) preso a: (*1*) um grupo amino (NH_3^+); (*2*) um grupo carboxila (COO^-); (*3*) um átomo de hidrogênio (*H*); e (*4*) um grupo radical, indicado como R. Nas estruturas dos 20 aminoácidos comuns, as partes em preto são comuns a todos os aminoácidos, e as partes em vermelho são os grupos R.

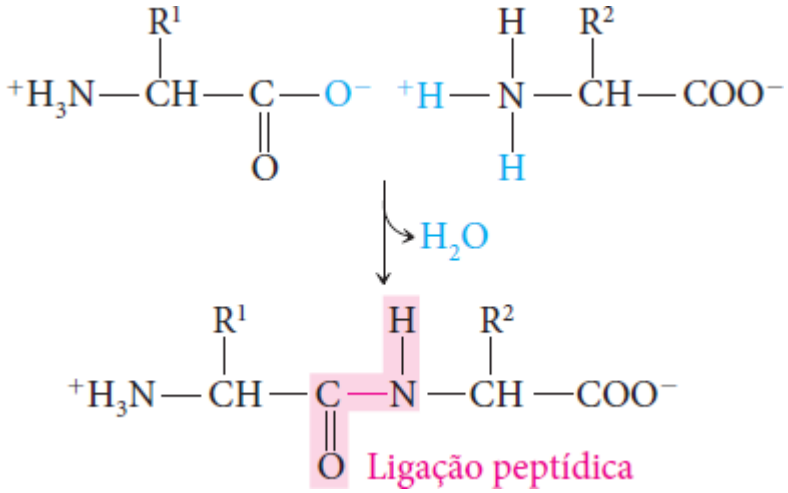


Figura 15.4 Os aminoácidos são unidos por ligações peptídicas. Em uma ligação peptídica (*vermelho*) o grupo carboxila de um aminoácido é preso por ligação covalente ao grupo amino do outro aminoácido.

Conceitos

Os produtos de muitos genes são proteínas cujas ações produzem os traços especificados por esses genes. As proteínas são polímeros com aminoácidos ligados por ligações peptídicas. A sequência de aminoácidos de uma proteína é sua estrutura primária. Essa estrutura se dobra para criar as estruturas secundárias e terciárias; duas ou mais cadeias polipeptídicas podem se associar para criar uma estrutura quaternária.

Checagem dos conceitos 2

O que determina basicamente as estruturas secundárias e terciárias de uma proteína?

15.2 O código genético determina como a sequência de nucleotídios especifica a sequência de aminoácidos de uma proteína

Em 1953, James Watson e Francis Crick esclareceram a estrutura do DNA e identificaram a sequência de bases como a carreadora da informação genética (ver *Capítulo 10*). Entretanto, a forma na qual a sequência de base do DNA especifica as sequências de aminoácidos das proteínas (o código genético) permaneceu indefinida por mais 10 anos.

Uma das primeiras perguntas a ser discutida sobre o código genético era quantos nucleotídios seriam necessários para especificar um único aminoácido. O conjunto de nucleotídios que codifica um único aminoácido – a unidade básica do código genético – é um códon (ver *Capítulo 14*). Muitos investigadores reconheceram que os códons precisam ter um mínimo de três nucleotídios. Cada posição do nucleotídio no mRNA pode ser ocupada por uma destas quatro bases: A, G, C ou U. Se um códon fosse composto por um único nucleotídio, apenas quatro códons diferentes (A, G, C e U) seriam possíveis, o que não é suficiente para codificar os 20 diferentes aminoácidos comumente encontrados nas proteínas. Se os códons fossem compostos por dois nucleotídios cada (p. ex., GU, AC etc.), existiriam 4 × 4 = 16 possíveis códons – ainda não seria suficiente para codificar todos os 20 aminoácidos. Com três nucleotídios por códon, há 4 × 4 × 4 = 64 possíveis códons, que é mais do que suficiente para especificar 20 diferentes aminoácidos. Portanto, um *código de trincas* com três nucleotídios por códon é a forma mais eficiente para codificar todos os 20 aminoácidos. Usando as mutações nos bacteriófagos, Francis Crick e seus colegas confirmaram, em 1961, que o código genético é na verdade um código de trincas. Resolva os Problemas 20 e 21

Conceitos

O código genético é constituído por trincas de nucelotídios (códon), no qual três nucleotídios codificam cada aminoácido em uma proteína.

Checagem dos conceitos 3

Um códon é:

a. Um dos três nucleotídios que codificam um aminoácido.
b. Três nucleotídios que codificam um aminoácido.
c. Três aminoácidos que codificam um nucleotídio.

d. Uma das quatro bases no DNA.

Como decifrar o código genético

Quando foi determinado que o código genético era composto por códons com três nucleotídios de comprimento, o próximo passo foi determinar quais grupos de três nucleotídios especificavam qual aminoácido. Pela lógica, a maneira mais fácil para decifrar o código seria determinar a sequência de base de um segmento de RNA, adicioná-lo a um tubo de teste contendo todos os componentes necessários para tradução e deixá-lo direcionar a síntese de uma proteína. A sequência de aminoácidos da proteína recém-sintetizada poderia ser determinada, e sua sequência seria comparada com a do RNA. Infelizmente, não existe uma forma para determinar a sequência de nucleotídios de um pedaço de RNA, então foram necessários métodos indiretos para decifrar o código.

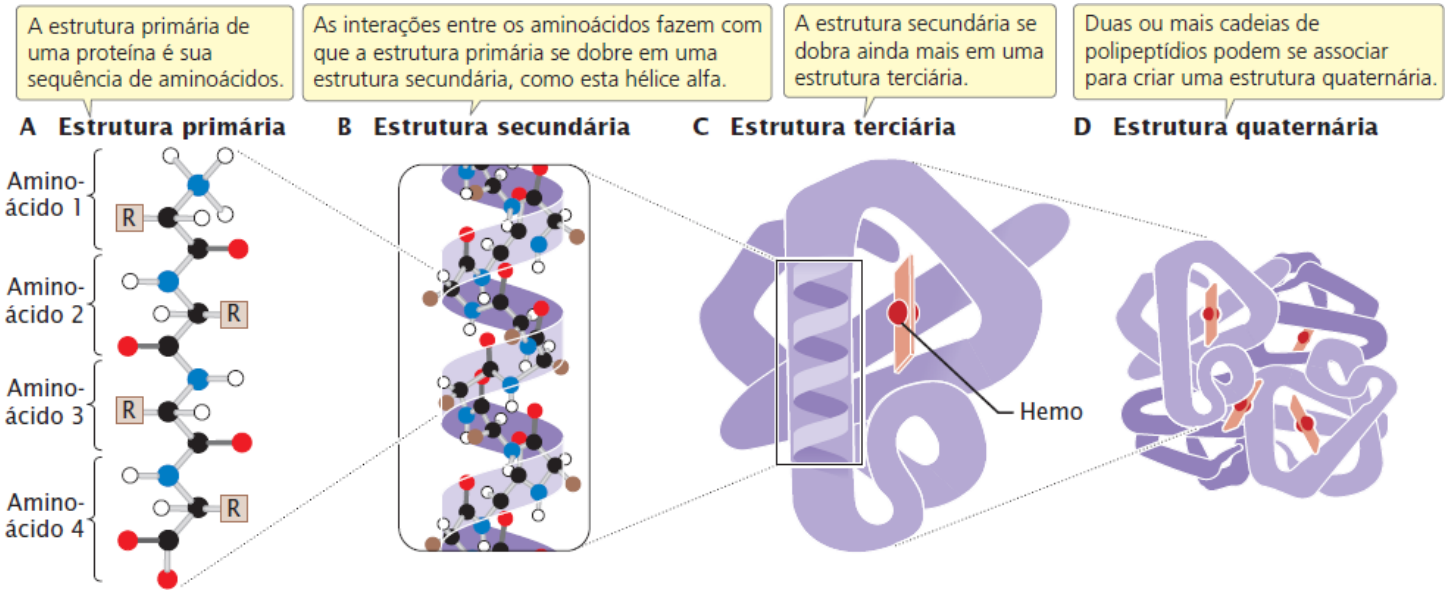
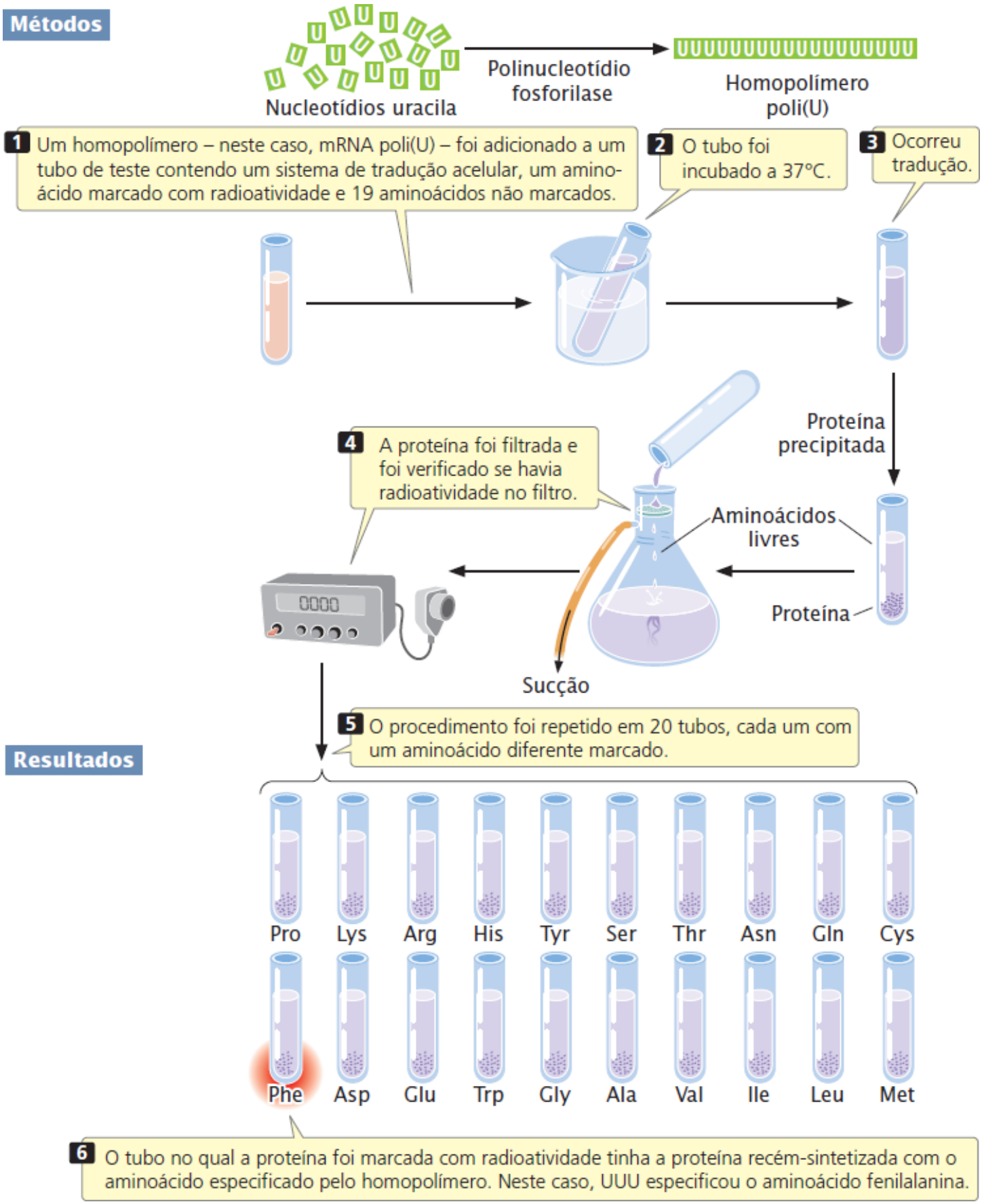


Figura 15.5 **As proteínas têm vários níveis de organização estrutural.**

Uso de homopolímeros. As primeiras pistas do código genético surgiram em 1961, a partir do trabalho de Marshall Nirenberg e Johann Heinrich Matthaei. Esses investigadores criaram RNAs sintéticos ao usar uma enzima chamada polinucleotídio fosforilase. Diferente da RNA polimerase, a polinucleotídio fosforilase não precisa de um molde; ela une aleatoriamente qualquer nucleotídio de RNA que esteja disponível. Os primeiros mRNAs sintéticos usados por Nirenberg e Matthaei eram homopolímeros, moléculas de RNA compostas por um único tipo de nucleotídio. Por exemplo, ao adicionar a polinucleotídio fosforilase a uma solução de nucleotídios uracila, eles geram moléculas de RNA compostas unicamente de nucleotídios uracila e, assim, têm apenas códons UUU (**Figura 15.6**). Esses poli(U) RNAs eram então adicionados a 20 tubos, cada um contendo os componentes necessários para tradução e todos os 20 aminoácidos. Um diferente aminoácido foi marcado com radioatividade em cada um dos 20 tubos. A proteína radioativa surgia em apenas um dos tubos – aquele contendo a fenilalanina marcada (ver Figura 15.6). Esse resultado mostrou que o códon UUU especifica o aminoácido fenilalanina. Os resultados de experimentos semelhantes usando RNA poli(C) e poli(A) demonstraram que o CCC codifica prolina e o AAA codifica lisina; por motivos técnicos, não foi possível interpretar os resultados do poli(G).

Experimento

Pergunta: Quais aminoácidos são especificados pelos códons compostos por apenas um tipo de base?



Conclusão: UUU codifica fenilalanina; em outros experimentos, AAA codificou lisina e CCC codificou prolina.

Figura 15.6 Nirenberg e Matthaei desenvolveram um método para identificar o aminoácido especificado por um homopolímero.

Uso de copolímeros aleatórios. Para obter informações sobre códons adicionais, Nirenberg e seus colegas criaram RNAs sintéticos com duas ou três bases diferentes. Como a polinucleotídio fosforilase incorpora os nucleotídios de maneira aleatória,

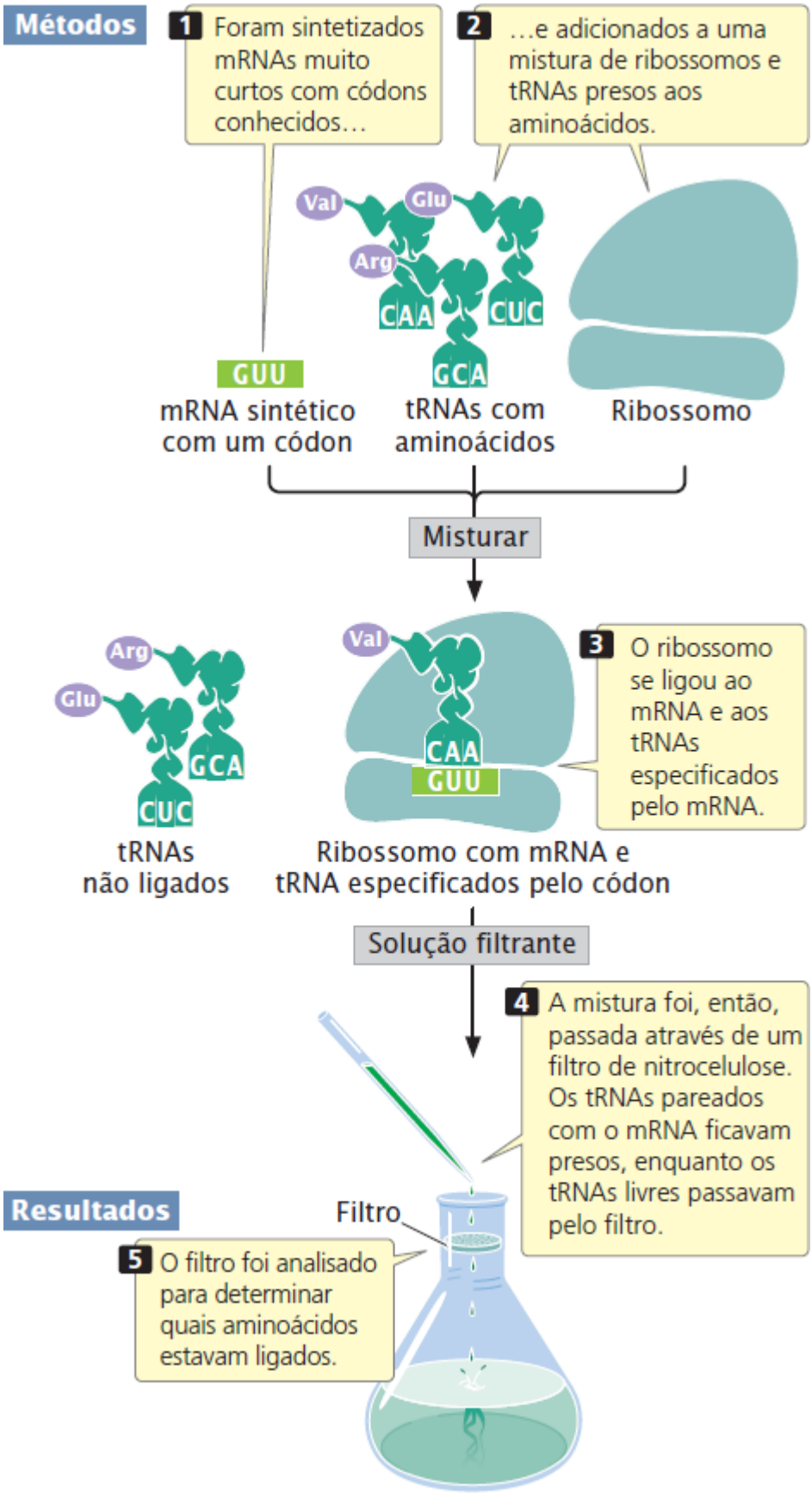
esses RNAs têm misturas aleatórias das bases e são chamados de copolímeros aleatórios. Por exemplo, quando os nucleotídios adenina e citosina eram misturados com a polinucleotídio fosforilase, as moléculas de RNA produziam oito tipos de códons diferentes: AAA, AAC, ACC, ACA, CAA, CCA, CAC e CCC. Esses RNAs poli(AC) produziam proteínas com seis tipos de aminoácidos diferentes: asparagina, glutamina, histidina, lisina, prolina e treonina.

As proporções de diferentes aminoácidos nas proteínas dependiam da razão dos dois nucleotídios usados para criar o mRNA sintético, e a probabilidade teórica de encontrar um códon particular poderia ser calculada a partir das razões das bases. Se uma razão 4:1 de C para A for usada para produzir o RNA, então a probabilidade de C estar em qualquer posição em um códon é 4/5, e a probabilidade de A estar nele é 1/5. Com a incorporação aleatória das bases, a probabilidade de qualquer um dos códons com dois Cs e um A (CCA, CAC ou ACC) deve ser 4/5 × 4/5 × 1/5 = 16/125 = 0,13, ou 13%, e a probabilidade de qualquer códon com dois As e um C (AAC, ACA ou CAA) deve ser 1/5 × 1/5 × 4/5 = 4/125 = 0,032, ou cerca de 3%. Portanto, um aminoácido codificado por dois Cs e um A será mais comum que um aminoácido codificado por dois As e um C. Ao comparar as porcentagens dos aminoácidos nas proteínas produzidas pelos copolímeros aleatórios com as frequências teóricas esperadas para os códons, Nirenberg e seus colegas poderiam deduzir informações sobre a *composição das bases* dos códons. Esses experimentos não revelaram nada sobre a *sequência* de bases do códon; a histidina era codificada por um códon com dois Cs e um A, mas não se sabia se o códon era ACC, CAC ou CCA. Havia outros problemas com este método: os cálculos teóricos dependiam da incorporação aleatória das bases, que nem sempre ocorria, e como o código genético é redundante, às vezes muitos códons diferentes especificam o mesmo aminoácido.

Uso de tRNAs ligados a ribossomo. Para superar as limitações dos copolímeros aleatórios, Nirenberg e Philip Leder desenvolveram outra técnica, em 1964, que usava tRNAs ligados a ribossomos. Eles descobriram que uma sequência muito curta de mRNA – até mesmo composta por um único códon – se ligaria a um ribossomo. O códon no mRNA curto teria o par de bases do anticódon pareado no RNA transportador que carregava o aminoácido especificado pelo códon (**Figura 15.7**). Os mRNAs curtos que eram ligados aos ribossomos eram misturados com tRNAs e aminoácidos e essa mistura era passada por um filtro de nitrocelulose. Os tRNAs que estavam pareados com o mRNA ligado ao ribossomo ficavam retidos no filtro, enquanto os tRNAs livres passavam pelo filtro. A vantagem deste sistema era que ele poderia ser usado com moléculas de mRNA sintéticas muito curtas que poderiam ser sintetizadas com uma sequência conhecida. Nirenberg e Leder sintetizaram mais de 50 mRNAs curtos com códons conhecidos e os adicionaram individualmente a uma mistura de ribossomos e tRNAs. Eles isolaram os tRNAs que estavam ligados ao mRNA e aos ribossomos e determinaram quais aminoácidos havia nos tRNAs ligados. Por exemplo, o RNA sintético com o códon GUU reteve um tRNA ao qual uma valina estava ligada, enquanto os RNAs com códons UGU e UUG, não. Usando este método, Nirenberg e seus colegas conseguiram determinar os aminoácidos codificados por mais de 50 códons.

Experimento

Pergunta: Com o uso de tRNAs, quais outras combinações entre códons e aminoácidos poderiam ser determinadas?



Conclusão: Quando um mRNA com GUU foi adicionado, os tRNAs no filtro estavam ligados à valina; portanto, o códon GUU especifica valina. Muitos outros códons foram determinados usando este método.

Figura 15.7 Nirenberg e Leder usaram tRNAs ligados ao ribossomo para obter informações adicionais sobre o código genético.

Outros experimentos forneceram informações adicionais sobre o código genético, e ele foi totalmente decifrado em 1968. Na seção a seguir, vamos examinar algumas dessas características do código, que é tão importante para a biologia moderna que

Francis Crick comparou seu lugar ao da tabela periódica dos elementos na química.

Degeneração do código

Um aminoácido é codificado por três nucleotídios consecutivos no mRNA, e cada nucleotídio pode ter uma das quatro bases possíveis (A, G, C e U) em cada posição de nucleotídio, resultando em 4^3 = 64 possíveis códons (**Figura 15.8**). Três desses códons são códons de parada, especificando o término da tradução. Assim, 61 códons, chamados **códons com sentido ou senso**, codificam os aminoácidos. Como existem 61 códons senso e apenas 20 diferentes aminoácidos comumente encontrados nas proteínas, o código tem mais informação que o necessário para especificar os aminoácidos, e por isso dizemos que é um **código genético degenerado**. Esta expressão não significa que o código genético está corrompido; *degenerado* é um termo que Francis Crick pegou emprestado da física quântica, que descreve múltiplos estados físicos com significado equivalente. A degeneração do código genético significa que os aminoácidos podem ser especificados por mais de um códon. Apenas o triptofano e a metionina são codificados por um único códon (ver Figura 15.8). Outros aminoácidos são especificados por dois códons, e alguns, como a leucina, são especificados por seis diferentes códons. Os códons que especificam o mesmo aminoácido são chamados de **códons sinônimos**, como as palavras sinônimas são diferentes palavras com o mesmo significado.

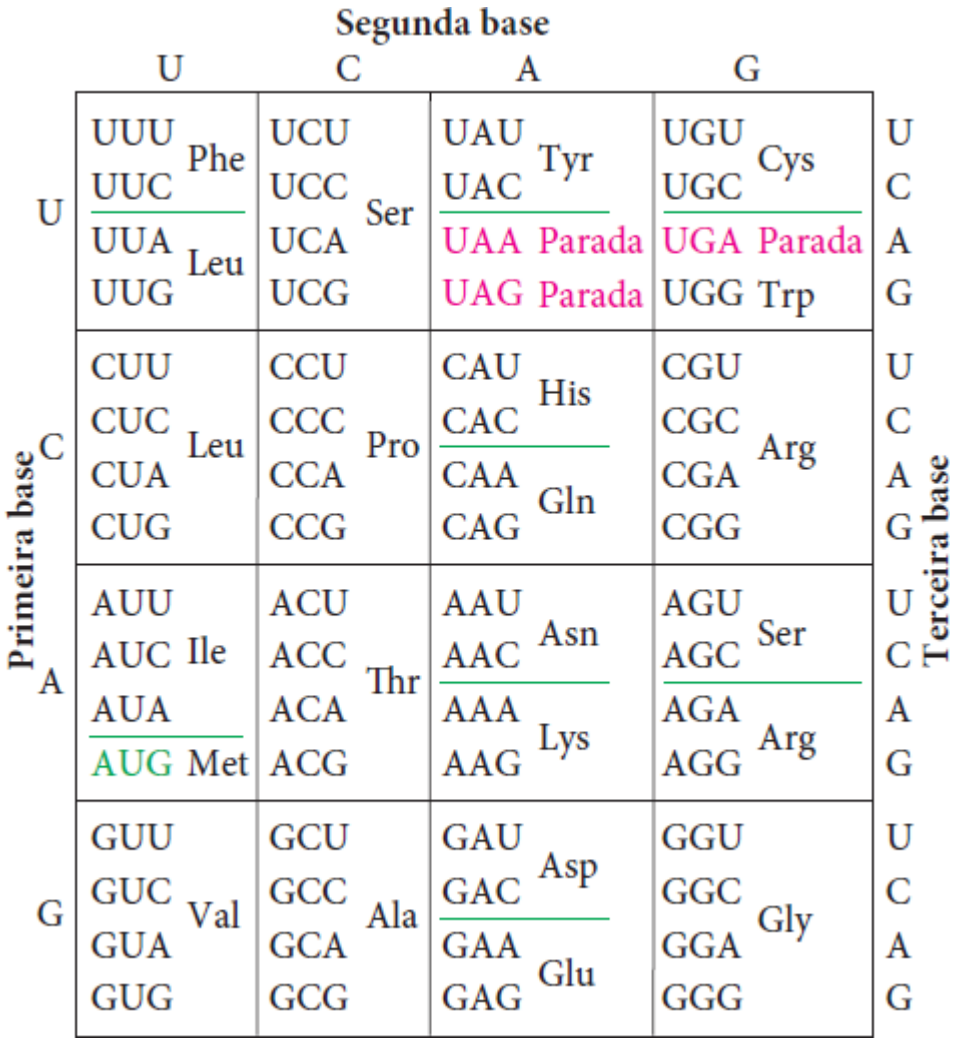


Figura 15.8 O código genético é composto por 64 códons. Os aminoácidos especificados por cada códon são indicados na abreviação de três letras. Os códons são escritos 5′ → 3′, como eles aparecem no mRNA. AUG é um códon de iniciação; UAA, UAG e UGA são códons de terminação (parada).

Como aprendemos no Capítulo 14, os tRNAs servem como moléculas adaptadoras, ligando-se a aminoácidos particulares e entregando-os a um ribossomo, onde os aminoácidos são então montados em cadeias de polipeptídios. Cada tipo de tRNA se prende a um único tipo de aminoácido. As células da maioria dos organismos têm cerca de 30 a 50 diferentes tRNAs, e ainda existem apenas 20 aminoácidos diferentes nas proteínas. Assim, alguns aminoácidos são carregados por mais de um tRNA. Diferentes tRNAs que aceitam o mesmo aminoácido, mas têm diferentes anticódons, são chamados de **tRNA isoaceptores**.

Mesmo que alguns aminoácidos tenham múltiplos tRNAs (isoaceptores), ainda existem mais códons do que anticódons, pois um anticódon pode parear com diferentes códons graças à flexibilidade no pareamento de base na terceira posição do códon. O exame da Figura 15.8 revela que muitos códons sinônimos são diferentes apenas na terceira posição. Por exemplo, a alanina é

codificada pelos códons GCU, GCC, GCA e GCG, todos começando com GC. Quando o códon no mRNA e o anticódon do tRNA se unem (**Figura 15.9**), a primeira base (5′) do códon pareia com a terceira base (3′) do anticódon, rigorosamente de acordo com as regras de Watson e Crick: A com U; C com G. A seguir, as bases do meio do códon e anticódon pareiam, também seguindo rigorosamente as regras de Watson e Crick. Após esses pares serem ligados por meio de pontes de hidrogênio, as terceiras bases pareiam fracamente e podem ser mais flexíveis, ou **oscilatórias** (***wobble***), no seu pareamento.

Em 1966, Francis Crick desenvolveu a hipótese da oscilação (*wobble*), segundo a qual alguns pareamentos de base não padrão poderiam ocorrer na terceira posição de um códon. Por exemplo, um G no anticódon pode parear com um C ou um U na terceira posição do códon (ver *Capítulo 14* e **Quadro 15.2**: observe que a inosina nesse quadro é uma das bases modificadas encontradas nos tRNAs). Algo importante a se lembrar sobre a oscilação (*wobble*) é que ela torna possível que alguns tRNAs pareiem com mais de um códon em um mRNA; assim, de 30 a 50 tRNAs podem parear com 61 códons senso. Alguns códons são sinônimos por meio da oscilação. Resolva o Problema 26

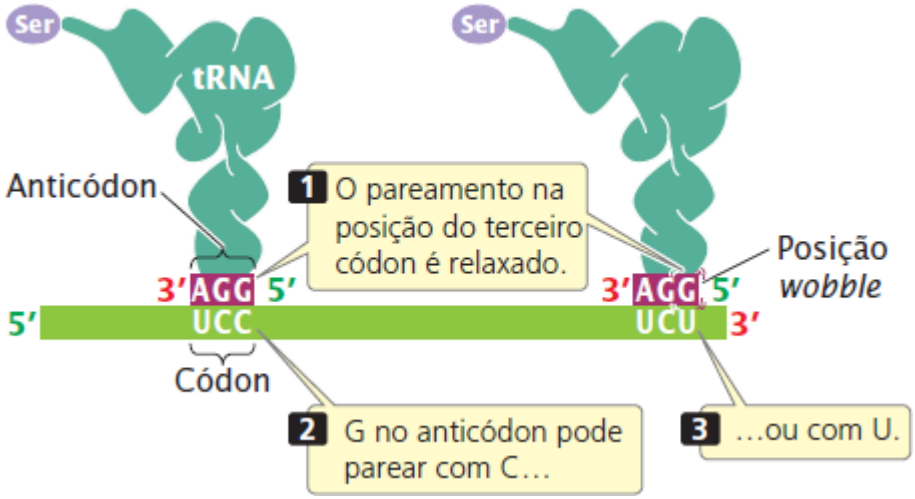
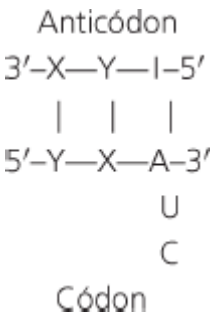


Figura 15.9 A oscilação (*wobble*) pode existir no pareamento de um códon e anticódon. O mRNA e o tRNA pareiam de maneira antiparalela. O pareamento na primeira e na segunda posições do códon está de acordo com as regras de pareamento de Watson e Crick (A com U, G com C); entretanto, as regras de pareamento são relaxadas na terceira posição do códon, e o G no anticódon pode parear com U ou C no códon neste exemplo.

Quadro 15.2 Regras de oscilação (*wobble*).†

Primeira posição do anticódon	Terceira posição do códon	Pareamento
C	G	Anticódon 3′–X—Y—C–5′ | | | 5′–Y—X—G–3′ Códon
G	U ou C	Anticódon 3′–X—Y—G–5′ | | | 5′–Y—X—U–3′ C Códon
A	U	Anticódon 3′–X—Y—A–5′ | | | 5′–Y—X—U–3′ Códon

U	A ou G	
I (inosina)*	A, U ou C	

†Essas regras indicam quais bases na terceira posição (extremidade 3′) do códon do mRNA podem parear com as bases na primeira posição (extremidade 5′) do anticódon do tRNA.
*Inosina é uma das bases modificadas encontradas nos tRNAs.

Conceitos

O código genético é composto por 61 códons senso que especificam os 20 aminoácidos comuns; o código é degenerado, o significa que alguns aminoácidos são codificados por mais de um códon. Os tRNAs isoaceptores são diferentes tRNAs com diferentes anticódons que especificam o mesmo aminoácido. Também, a oscilação na terceira posição do códon torna possível que diferentes códons especifiquem o mesmo aminoácido.

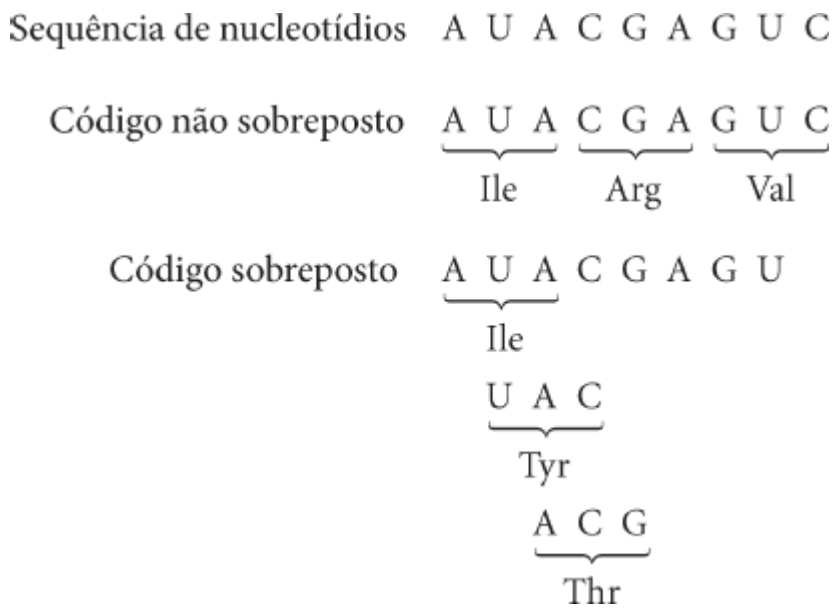
Checagem dos conceitos 4

Por meio da oscilação, um único _____ pode parear com mais de um _____.

a. Códon, anticódon.
b. Grupo de três nucleotídios no DNA, códon no mRNA.
c. tRNA, aminoácido.
d. Anticódon, códon.

Quadro de leitura e códons de iniciação

Achados de estudos iniciais do código genético indicaram que em geral o **código genético não é sobreposto**. Um código sobreposto é um no qual um único nucleotídio pode ser incluído em mais de um códon, assim:



Entretanto, cada nucleotídio é parte de um único códon. Alguns genes sobrepostos são encontrados nos vírus, mas os códons dentro do mesmo gene não se sobrepõem e o código é considerado como não sobreposto.

Para qualquer sequência de nucleotídios, existem três potenciais conjuntos de códons – três formas nas quais a sequência pode ser lida em grupos de três. Cada maneira diferente de ler a sequência é chamada de **quadro de leitura**, e qualquer sequência de nucleotídios tem três quadros de leitura potenciais. Os três quadros de leituras têm conjuntos completamente diferentes de códons e, portanto, especificarão proteínas com sequências de aminoácidos muito diferentes. Assim, é essencial que o maquinário de tradução use o quadro de leitura correto. Como o quadro de leitura correto é estabelecido? O quadro leitura é definido pelo **códon de iniciação**, que é o primeiro códon do mRNA a especificar um aminoácido. Após o códon de iniciação, outros códons são lidos como grupos sucessivos de três nucleotídios. Nenhuma base é pulada entre os códons, então não há sinais de pontuação separando os códons.

O códon de iniciação em geral é AUG, embora GUG e UUG sejam usados em raros casos. O códon de iniciação não é apenas uma sequência que marca o início da tradução; ele especifica um aminoácido. Nas células bacterianas, o primeiro AUG codifica um tipo modificado de metionina, *N*-formilmetionina; todas as proteínas nas bactérias inicialmente começam com esse aminoácido, mas o grupo formila (ou em alguns casos, o aminoácido inteiro) pode ser removido após a síntese da proteína. Quando o códon AUG está em uma posição interna em um gene, ele codifica a metionina não formilada. Nas células de archaea e eucarióticas, AUG especifica a metionina não formilada tanto na posição de início quanto nas posições internas. Nas bactérias e nos eucariotos existem diferentes tRNAs para a metionina iniciadora (designada $tRNA_i^{fMet}$ nas bactérias e $tRNA_i^{Met}$ nos eucariotos) e metionina interna (designada $tRNA^{Met}$).

Códons de terminação

Três códons – UAA, UAG e UGA – não codificam aminoácidos. Esses códons sinalizam o final da proteína nas células bacterianas e eucarióticas e são chamados de **códons de parada**, **códons de terminação** ou **códons *nonsense*** ou **sem sentido**. Nenhuma molécula de tRNA tem anticódons que pareiam com códons de terminação.

Universalidade do código

Por muitos anos, acreditava-se que o **código genético** fosse **universal**, o que significa que cada códon especifica o mesmo aminoácido em todos os organismos. Agora sabemos que o código genético é quase, mas não completamente, universal; algumas exceções foram encontradas. A maioria dessas exceções são os códons de terminação, mas existem alguns casos nos quais um códon senso substitui outro. A maioria das exceções é encontrada nos genes mitocondriais; alguns códons não universais também foram detectados nos genes nucleares do DNA de protozoários e bactérias (**Quadro 15.3**). Resolva o Problema 22

Conceitos

Cada sequência de nucleotídios tem três potenciais quadros de leitura. O quadro de leitura correto é determinado pelo códon de iniciação. O final de uma sequência codificadora de proteína é marcado por um códon de terminação. Com algumas exceções, todos os organismos usam o mesmo código genético.

Checagem dos conceitos 5

Os códons de iniciação e terminação especificam um aminoácido? Se sim, qual deles?

Conceitos conectantes

Características do código genético

Agora analisaremos várias características do código genético. Revisaremos essas características:

1. O código genético consiste em uma sequência de nucleotídios no DNA ou RNA. Há quatro letras no código, correspondendo a quatro bases: A, G, C e U (T no DNA).
2. O código genético é um código de trincas. Cada aminoácido é codificado por uma sequência de três nucleotídios consecutivos, chamada códon.
3. O código genético é degenerado; de 64 códons, 61 codificam apenas 20 aminoácidos nas proteínas (3 são códons de terminação). Alguns códons são sinônimos, especificando o mesmo aminoácido.
4. Os tRNA isoaceptores são tRNAs com diferentes anticódons que aceitam o mesmo aminoácido; a oscilação (*wobble*) possibilita que o anticódon em um tipo de tRNA pareie com mais de um tipo de códon no mRNA.
5. O código em geral não se sobrepõe; cada nucleotídio em uma sequência de mRNA pertence a um único quadro de leitura.
6. O quadro de leitura é definido por um códon de iniciação, em geral AUG.
7. Quando um quadro de leitura é definido, os códons são lidos como grupos sucessivos de três nucleotídios.
8. Qualquer um dos três códons de terminação (UAA, UAG e UGA) pode sinalizar o final de uma proteína; nenhum aminoácido é codificado pelos códons de terminação.
9. O código é quase universal.

15.3 Aminoácidos são montados em uma proteína por meio da tradução

Agora que estamos familiarizados com o código genético, podemos começar a estudar como os aminoácidos são montados em proteínas. Como temos mais informações sobre a tradução nas bactérias, vamos nos concentrar principalmente na tradução bacteriana. Em muitos aspectos, a tradução eucariótica é semelhante, embora existam algumas diferenças importantes.

Quadro 15.3 Algumas exceções para o código genético universal.

Genoma	Códon	Código universal	Código alterado
DNA bacteriano			
Mycoplasma capricolum	UGA	Parada	Trp
DNA mitocondrial			
Humano	UGA	Parada	Trp
Humano	AUA	Ile	Met
Humano	AGA, AGG	Arg	Parada
Levedura	UGA	Parada	Trp
Tripanossomos	UGA	Parada	Trp
Plantas	CGG	Arg	Trp
DNA nuclear			
Tetrahymena	UAA	Parada	Gln
Paramecium	UAG	Parada	Gln

Lembre-se de que apenas os mRNAs são traduzidos em proteínas. A tradução ocorre nos ribossomos; de fato, os ribossomos podem ser considerados como maquinários móveis de síntese de proteínas. Por meio de várias técnicas, foi produzida uma visão detalhada da estrutura do ribossomo nos últimos anos, que melhorou em muito nossa compreensão sobre a tradução. Um ribossomo se prende próximo à extremidade 5′ da fita de mRNA e se move para a extremidade 3′, traduzindo o códon à medida que se desloca (**Figura 15.10**). A síntese começa na extremidade amino da proteína, e a proteína é alongada com a adição de novos aminoácidos à extremidade carboxila. A síntese de proteínas inclui uma série de interações RNA-RNA: interações entre mRNA e o rRNA que segura o mRNA no ribossomo, entre o códon no mRNA e o anticódon no tRNA e entre o tRNA e os rRNAs do ribossomo.

A síntese de proteína pode ser dividida em quatro estágios: (1) carga do tRNA, no qual o tRNA se liga aos aminoácidos; (2) início, no qual os componentes necessários para a tradução são montados no ribossomo; (3) alongamento, no qual os aminoácidos são unidos, um por vez, à cadeia de polipeptídios crescente; e (4) término, no qual a síntese de proteína para no códon de terminação e os componentes da tradução são liberados do ribossomo.

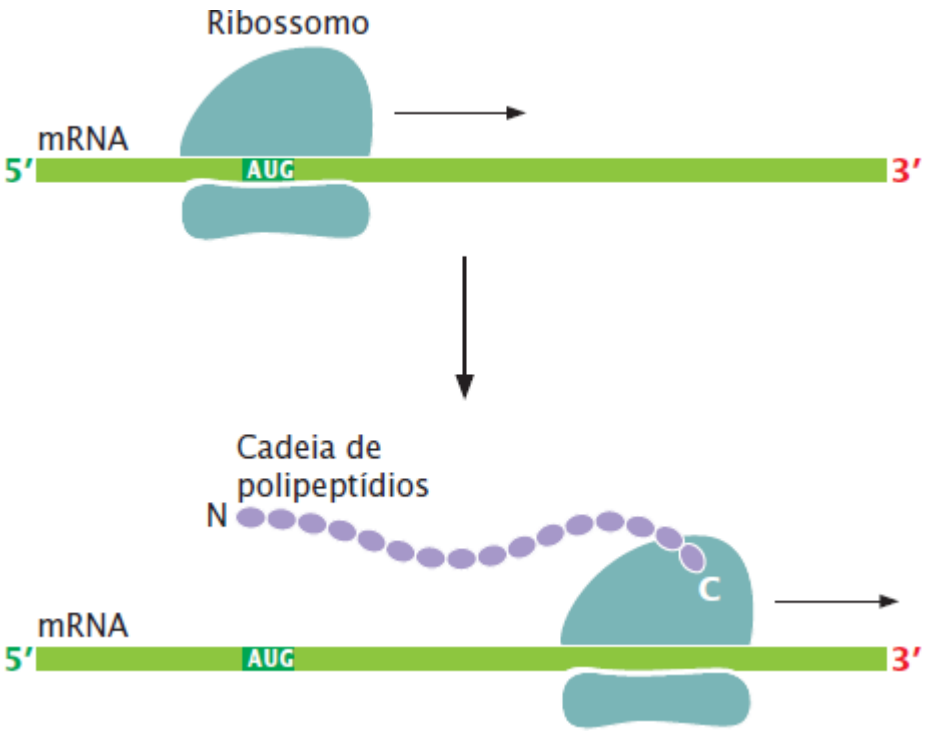


Figura 15.10 A tradução de uma molécula de mRNA ocorre em um ribossomo. A letra N representa a extremidade amino da proteína; C representa a extremidade carboxila.

Ligação dos aminoácidos aos RNAs transportadores

O primeiro estágio da tradução é a ligação das moléculas de tRNA a seus aminoácidos adequados, chamado tRNA carregado. Cada tRNA é específico para um aminoácido particular. Todos os tRNAs têm a sequência CCA na extremidade 3′, e o grupo carboxila (COO^-) do aminoácido é ligado ao nucleotídio adenina na extremidade 3′ do tRNA (**Figura 15.11**). Se cada tRNA é específico para determinado aminoácido, mas todos os aminoácidos estão ligados ao mesmo nucleotídio (A) na extremidade 3′ do tRNA, como o tRNA se liga a seu aminoácido adequado?

A chave para a especificidade entre um aminoácido e seu tRNA é um conjunto de enzimas chamadas de **aminoacil-tRNA sintetases**. Uma célula tem 20 diferentes aminoacil-tRNA sintetases, uma para cada um dos 20 aminoácidos. Cada sintetase reconhece um aminoácido particular, assim como todos os tRNAs que aceitam este aminoácido. O reconhecimento do aminoácido adequado por uma sintetase baseia-se em diferentes tamanhos, cargas e grupos R dos aminoácidos. O reconhecimento dos tRNAs por uma sintetase depende das diferentes sequências de nucleotídios dos tRNAs. Os pesquisadores identificaram quais nucleotídios são importantes no reconhecimento ao alterar diferentes nucleotídios em um tRNA particular e determinar se o tRNA alterado ainda é reconhecido por sua sintetase (**Figura 15.12**).

A ligação de um tRNA a seu aminoácido adequado, chamado **tRNA carregado**, requer energia, que é fornecida pelo trifosfato de adenosina (ATP):

$$\text{aminoácido} + \text{tRNA} + \text{ATP} \rightarrow \text{aminoacil-tRNA} + \text{AMP} + PP_i$$

Esta reação ocorre em duas etapas (**Figura 15.13**). Para identificar o tRNA aminoacetilado resultante, escrevemos a abreviação de três letras para o aminoácido na frente do tRNA; por exemplo, o aminoácido alanina (Ala) se fixa ao seu tRNA ($tRNA^{Ala}$), dando origem a seu aminoacil-tRNA ($Ala\text{-}tRNA^{Ala}$).

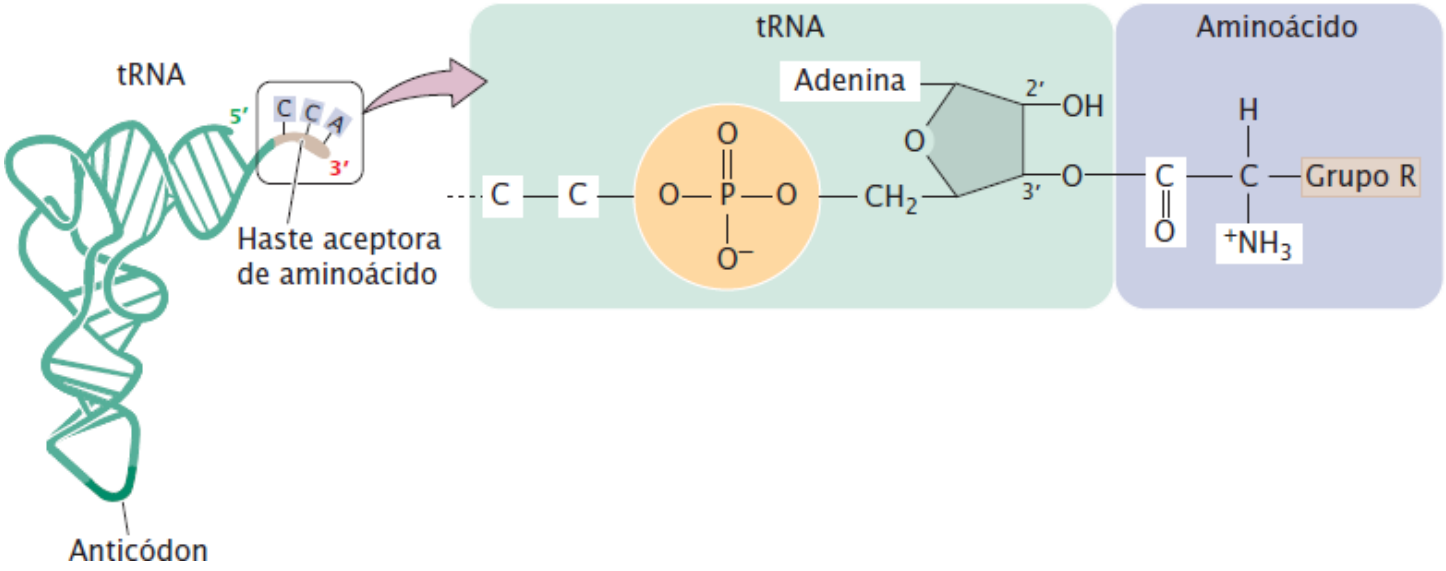


Figura 15.11 Um aminoácido se liga à extremidade 3′ de um tRNA. O grupo carboxila (COO^-) do aminoácido se liga ao grupo hidroxila do átomo de carbono 2′ ou 3′ do nucleotídio final na extremidade 3′ do tRNA, no qual a base é sempre adenina.

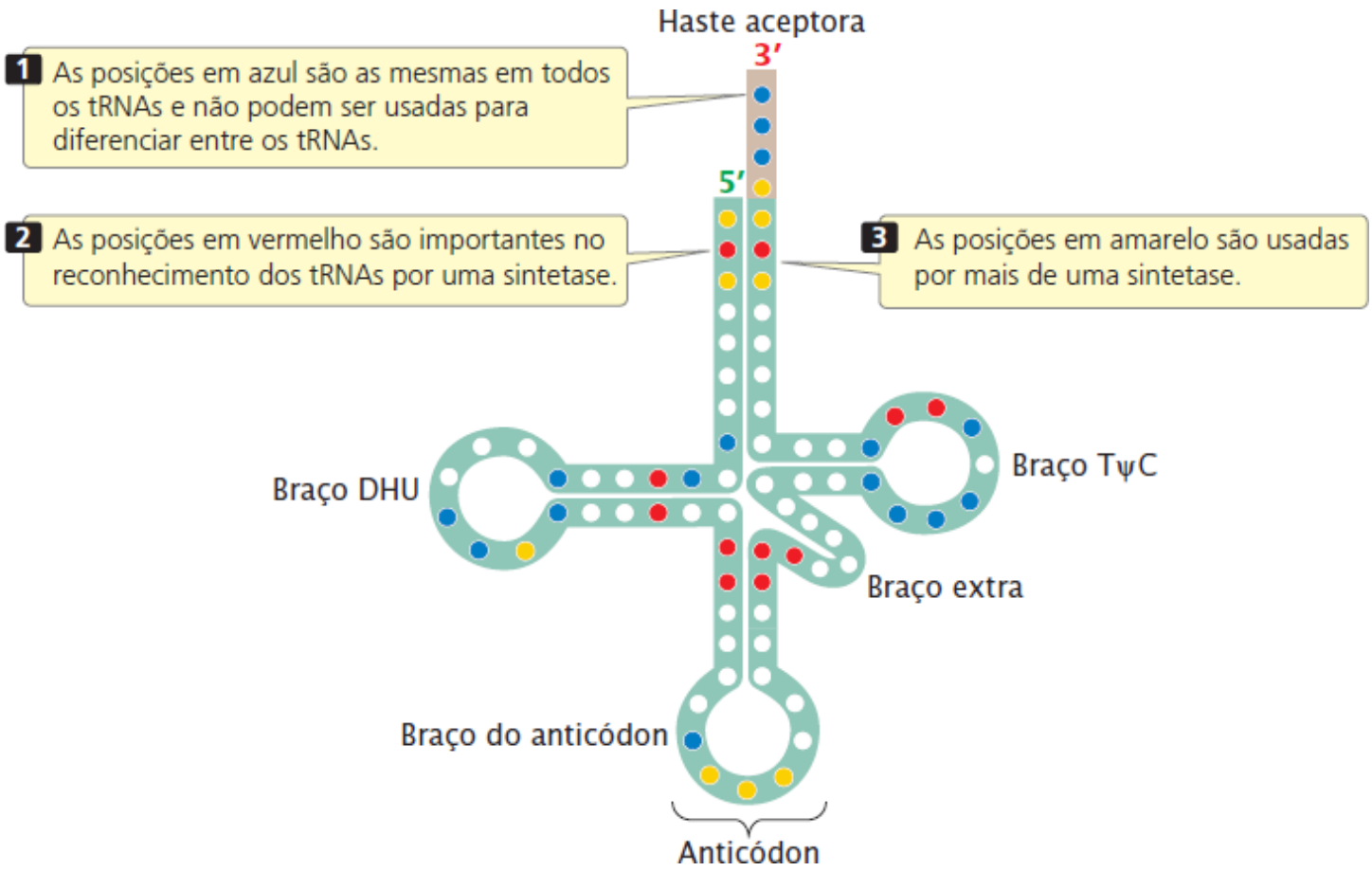


Figura 15.12 Algumas posições nas moléculas de tRNA são reconhecidas pela aminoacil-tRNA sintetase adequada.

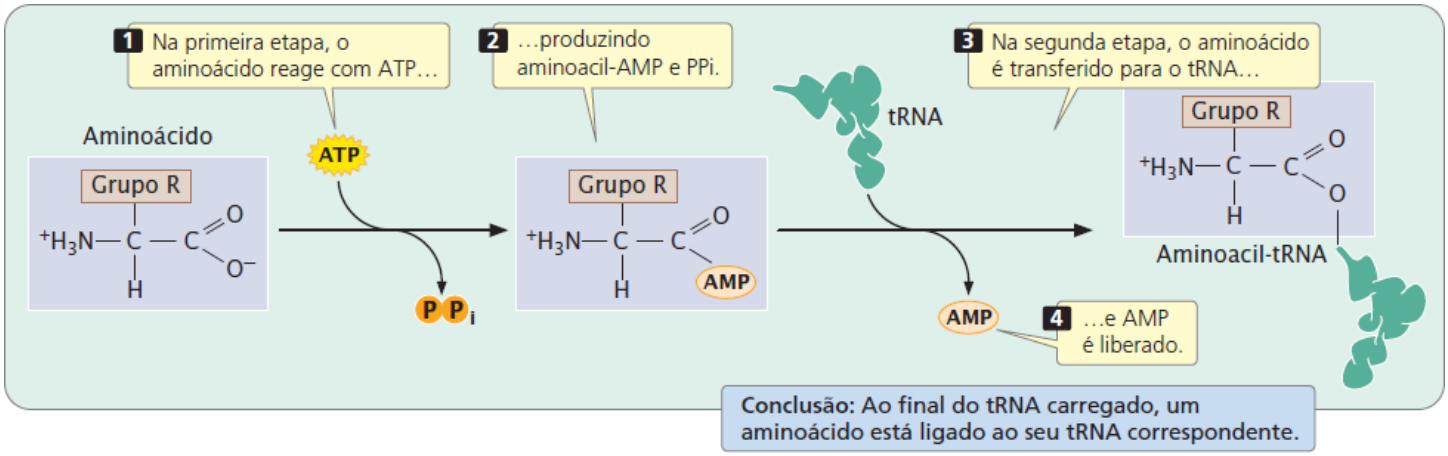


Figura 15.13 Um aminoácido se liga ao tRNA adequado por meio de uma reação de duas etapas.

Os erros no tRNA carregado ou aminoacilado são raros: eles ocorrem em apenas 1 em 10.000 ou 1 em 100.000 reações. Esta fidelidade deve-se em parte à edição (revisão) em muitas das sintetases. A edição detecta e remove aminoácidos incorretamente pareados dos tRNAs. Alguns agentes químicos antifúngicos trabalham ao prender os tRNAs no sítio de edição da enzima, evitando sua liberação e inibindo o processo de tradução nos fungos.

Conceitos

Os aminoácidos são ligados a tRNAs específicos por aminoacil-tRNA sintetases em uma reação de duas etapas que requer ATP.

Checagem dos conceitos 6

Os aminoácidos se ligam a que parte do tRNA?

a. Anticódon.
b. Braço DHU.
c. Extremidade 3′.
d. Extremidade 5′.

Iniciação da tradução

O segundo estágio no processo de síntese de proteínas é a iniciação. Nesse estágio, todos os componentes necessários para a síntese proteica são montados: (1) mRNA; (2) as subunidades menor e maior do ribossomo; (3) um conjunto de três proteínas chamadas de fatores de iniciação; (4) tRNA iniciador com a *N*-formilmetionina presa (fMet-tRNA^{fMet}); e (5) trifosfato de guanosina (GTP). O início inclui três etapas. Primeiro, o mRNA se liga à subunidade menor do ribossomo. Segundo, o tRNA iniciador se liga ao mRNA por meio do pareamento de base entre o códon e o anticódon. Terceiro, o ribossomo maior se une ao complexo de início. Agora vamos olhar cada uma dessas etapas com mais detalhes.

Iniciação nas bactérias. O ribossomo funcional das bactérias existe como duas subunidades, a subunidade menor 30S e a subunidade maior 50S (**Figura 15.14 A**). Uma molécula de mRNA pode se ligar à subunidade menor do ribossomo apenas quando as subunidades estão separadas. O **fator de iniciação 3** (IF-3) se liga à subunidade menor do ribossomo e evita que a subunidade maior se ligue durante o início (**Figura 15.14 B**). Outro fator, o **fator de iniciação 1** (IF-1), acentua a dissociação das subunidades maior e menor do ribossomo.

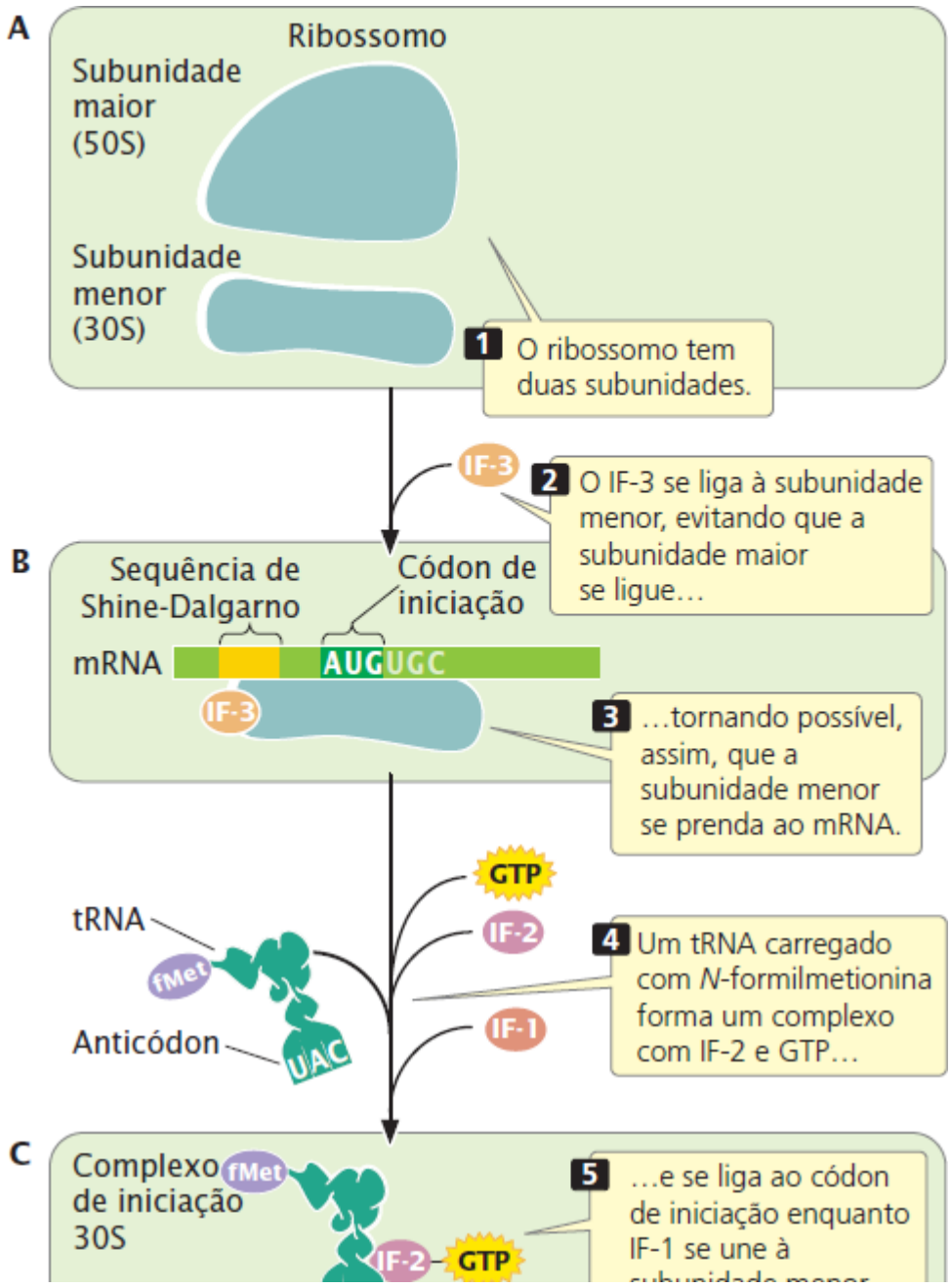
Onde o ribossomo se liga ao mRNA durante a iniciação da tradução? As sequências-chave no mRNA necessárias para a ligação ao ribossomo foram identificadas em experimentos projetados para possibilitar que o ribossomo se ligue ao mRNA, mas que não prossiga com a síntese de proteínas; o ribossomo é parado no sítio de início onde ocorre a ligação. Uma ribonuclease, que degrada todos o mRNA exceto a região coberta pelo ribossomo, é adicionada. O mRNA intacto pode ser separado do ribossomo e estudado. A sequência coberta pelo ribossomo durante o início tem de 30 a 40 nucleotídios e inclui o códon de iniciação AUG. No sítio de ligação do ribossomo está a sequência consenso de Shine-Dalgarno (**Figura 15.15**, ver também *Capítulo 14*), que é complementar a uma sequência de nucleotídios na extremidade 3′ do 16S rRNA (parte da subunidade menor do ribossomo). Durante o início, os nucleotídios na sequência de Shine-Dalgarno pareiam com seus nucleotídios complementares no 16S rRNA, fazendo com que a subunidade menor do ribossomo se prenda ao mRNA e posicione o ribossomo diretamente sobre o códon de iniciação. Essas sequências ligantes de ribossomo estão na região 5′ não traduzida do

mRNA.

O tRNA iniciador, fMet-tRNA^{fMet}, prende-se ao códon de iniciação (ver Figura 15.14 C). Essa etapa requer o **fator de iniciação 2** (IF-2), que forma um complexo com GTP.

Nesse ponto, o complexo de iniciação é composto por (1) a subunidade menor do ribossomo; (2) o mRNA; (3) o tRNA iniciador com seu aminoácido (fMet-tRNA^{fMet}); (4) uma molécula de GTP; e (5) vários fatores de iniciação. Esses componentes são coletivamente conhecidos como o **complexo de iniciação 30S** (ver Figura 15.14 C). No final da etapa de iniciação, IF-3 se dissocia da subunidade menor, fazendo com que a subunidade maior do ribossomo se una ao complexo de iniciação. A molécula de GTP (fornecida pelo IF-2) é hidrolisada a difosfato de guanosina (GDP), e os fatores de iniciação se dissociam (ver Figura 15.14 D). Quando a subunidade maior se unir ao complexo de iniciação, o complexo passa a ser chamado de **complexo de iniciação 70S**.

Iniciação nos eucariotos. Ocorrem eventos semelhantes na iniciação da tradução nas células eucarióticas, mas existem diferenças importantes. Nas células bacterianas, as sequências no 16S rRNA da subunidade menor do ribossomo se ligam à sequência de Shine-Dalgarno no mRNA. Não existe sequência consenso análoga no mRNA eucariótico. Pelo contrário, o cap da extremidade 5′ do mRNA eucariótico tem um papel crítico na iniciação da tradução. Em uma série de etapas, a subunidade menor do ribossomo eucariótico, os fatores de iniciação e o tRNA iniciador com seu aminoácido (Met-tRNA$_i^{Met}$) formam um complexo de iniciação que reconhece o cap e se liga a ele. O complexo de iniciação, então, se move ao longo (varre) o mRNA até localizar o primeiro códon AUG. A identificação do códon de iniciação é facilitada por uma sequência consenso (chamada de sequência Kozak) que circunda o códon de iniciação:



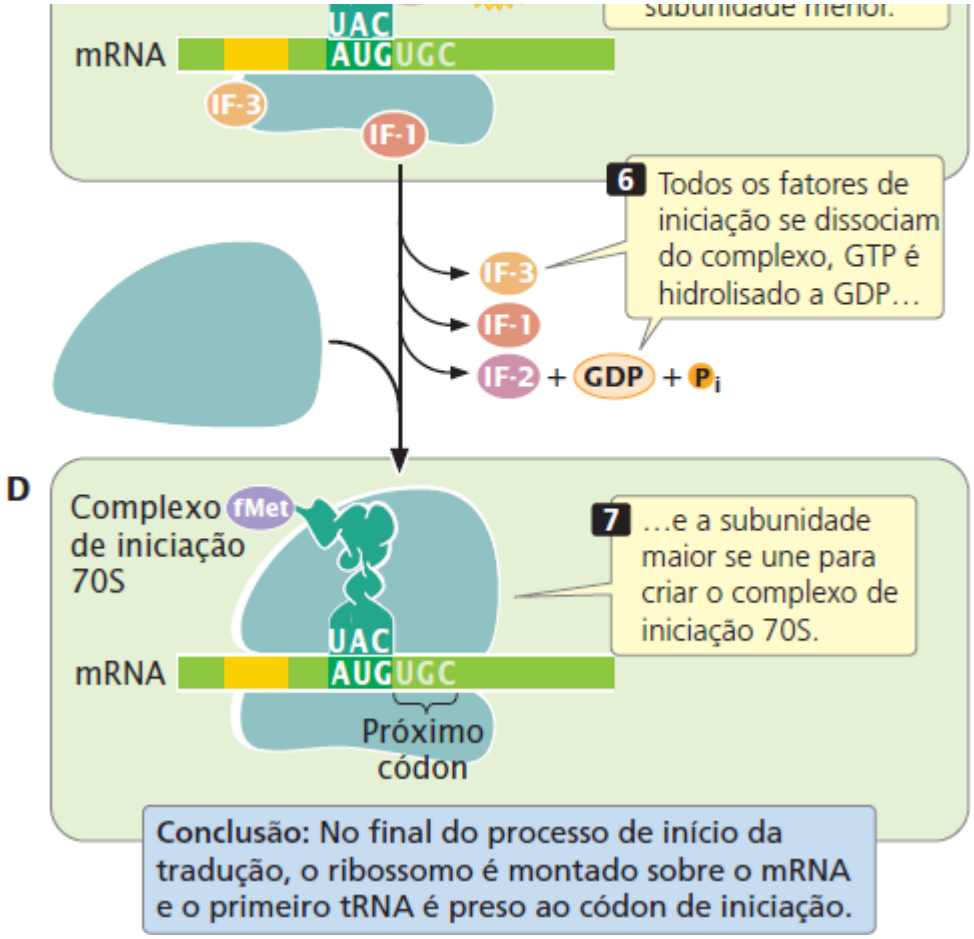


Figura 15.14 **A iniciação da tradução requer vários fatores de iniciação e GTP.**

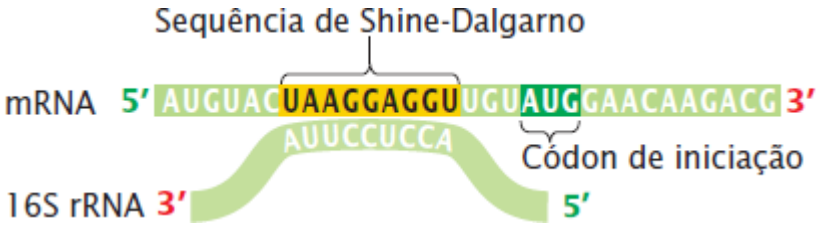
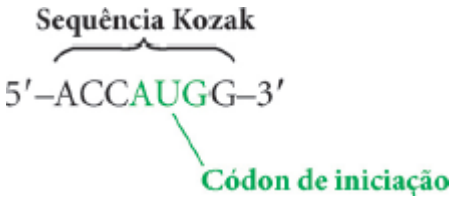


Figura 15.15 **A sequência consenso de Shine-Dalgarno no mRNA é necessária para a fixação da subunidade menor do ribossomo.**



Outra diferença importante é que a iniciação em eucariotos requer pelo menos sete fatores de iniciação. Alguns fatores mantêm as subunidades separadas, como o IF-3 faz nas células bacterianas. Outros reconhecem o cap 5′ no mRNA e fazem com que a subunidade menor do ribossomo se ligue a este. Outros ainda têm atividade da RNA helicase, que é usada para desenrolar estruturas secundárias que podem existir na região 5′ não traduzida do mRNA, possibilitando que a subunidade menor se desloque no mRNA até o códon de iniciação ser alcançado. Outros fatores de iniciação ajudam a trazer o Met-tRNA$_i^{Met}$ ao complexo de iniciação.

Nos eucariotos, o cap é ligado inicialmente a várias proteínas, uma das quais é o **complexo de ligação ao cap (CBC)**. O CBC ajuda a exportar o mRNA do núcleo e então promove o ciclo “pioneiro” ou inicial da tradução no citoplasma. Esse primeiro ciclo de tradução é importante na verificação de erros no mRNA (ver Vigilância do RNA mensageiro na próxima seção). Após o ciclo pioneiro de tradução, o CBC é substituído pelo fator de iniciação eucariótico 4E (eIF-4E), que promove a tradução contínua do mRNA.

A cauda poli(A) na extremidade 3′ do mRNA eucariótico também participa na iniciação da tradução. Durante o início, as proteínas que se fixam à cauda poli(A) interagem com proteínas que se ligam ao cap 5′, acentuando a ligação da subunidade

menor do ribossomo para a extremidade 5′ do mRNA. Essa interação indica que a extremidade 3′ do mRNA se dobra e se associa com o cap 5′ durante a iniciação da tradução, formando uma estrutura circular conhecida como alça fechada (**Figura 15.16**). Poucos mRNAs eucarióticos têm sítios de entrada interna do ribossomo, onde os ribossomos podem se ligar diretamente sem primeiro se ligar ao cap 5′.

Conceitos

Na iniciação da tradução nas células bacterianas, a subunidade menor do ribossomo se fixa ao mRNA, e um tRNA iniciador se fixa ao códon de iniciação. Esse processo requer vários fatores de iniciação (IF-1, IF-2 e IF-3) e GTP. Na etapa final, a subunidade maior do ribossomo se une ao complexo de iniciação.

Checagem dos conceitos 7

Durante o início da tradução, o ribossomo menor se liga a qual sequência consenso nas bactérias?

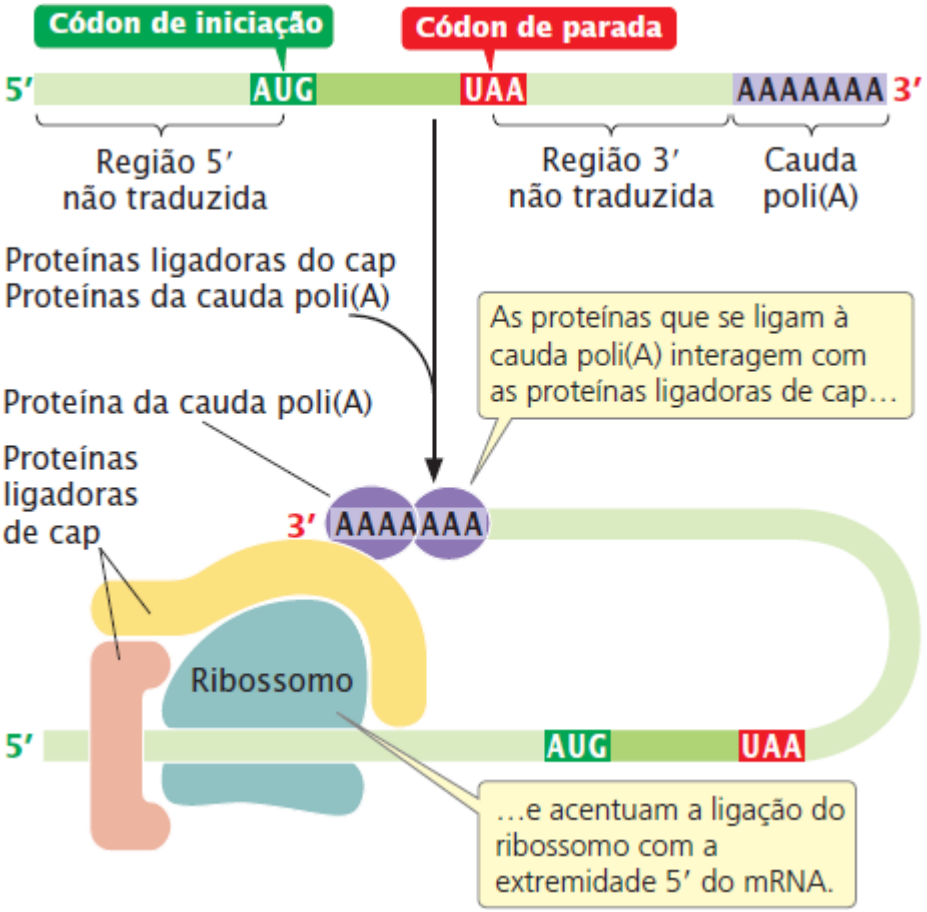


Figura 15.16 A cauda poli(A) do mRNA eucariótico participa da iniciação da tradução.

Alongamento

O próximo estágio na síntese de proteínas é o alongamento, no qual os aminoácidos são unidos para criar uma cadeia de polipeptídios. O alongamento requer (1) o complexo 70S já descrito; (2) tRNAs carregados com seus aminoácidos; (3) vários fatores de alongamento; e (4) GTP.

Um ribossomo tem três sítios que podem ser ocupados por tRNAs; o **sítio aminoacil** (ou **sítio A)**, o **sítio de peptidil** (ou **sítio P)** e o **sítio de saída** (ou **sítio E)** (**Figura 15.17 A**). O tRNA iniciador ocupa imediatamente o sítio P (o único sítio ao qual o fMet-tRNA^{fMet} é capaz de se ligar), mas todos os outros tRNAs entram primeiro no sítio A. No final da iniciação, o ribossomo está ligado ao mRNA e o fMet-tRNA^{fMet} é posicionado sobre o códon de iniciação AUG no sítio P; o sítio A adjacente está desocupado (ver Figura 15.17 A).

O alongamento ocorre em três etapas. Na primeira (**Figura 15.17 B**), um tRNA carregado se liga ao sítio A. A ligação ocorre quando o **fator Tu de alongamento** (EF-Tu) se une ao GTP e então a um tRNA carregado para formar um complexo de três partes. Esse complexo entra no sítio A do ribossomo, onde o anticódon no tRNA pareia com o códon no mRNA. Após o tRNA carregado entrar no sítio A, o GTP é rompido para GDP, e o complexo EF-Tu-GDP é liberado (**Figura 15.17 C**). O **fator Ts de**

alongamento (EF-Ts) regenera o EF-Tu-GDP para EF-Tu-GTP. Nas células eucarióticas, um conjunto semelhante de reações libera o tRNA carregado para o sítio A.

A segunda etapa do alongamento é a formação de uma ligação peptídica entre os aminoácidos que são presos aos tRNAs nos sítios P e A (**Figura 15.17 D**). A formação dessa ligação peptídica libera o aminoácido de seu tRNA no sítio P. A formação da ligação peptídica ocorre dentro do centro da peptidiltransferase, que é parte da subunidade maior do ribossomo. Evidências indicam que a atividade catalítica é uma propriedade do RNA ribossômico na subunidade maior do ribossomo (o 23S rRNA nas bactérias, o 28S RNA nos eucariotos); esse rRNA atua como uma ribozima (ver *Capítulo 13*).

A terceira etapa no alongamento é a **translocação** (**Figura 15.17 E**), o movimento do ribossomo no mRNA no sentido 5′ → 3′. Essa etapa posiciona o ribossomo sobre o próximo códon e requer o **fator G de alongamento** (EF-G) e a hidrólise de GTP para GDP. Como os tRNAs nos sítios P e A ainda estão ligados ao mRNA por meio do pareamento códon-anticódon, eles não se movem com o ribossomo à medida que ele transloca. Consequentemente, o ribossomo se desloca de modo que o tRNA que ocupava o sítio P agora ocupe o sítio E, a partir do qual ele se move para o citoplasma onde pode ser recarregado com outro aminoácido. A translocação também faz com que o tRNA que ocupava o sítio A (que está preso à cadeia de polipeptídios crescente) vá para o sítio P, deixando o sítio A aberto. Assim, o progresso de cada tRNA por meio do ribossomo durante o alongamento pode ser resumido como se segue: citoplasma → sítio A → sítio P → sítio E → citoplasma. Como já foi afirmado, o tRNA iniciador é uma exceção: ele se prende diretamente ao sítio P e nunca ocupa o sítio A.

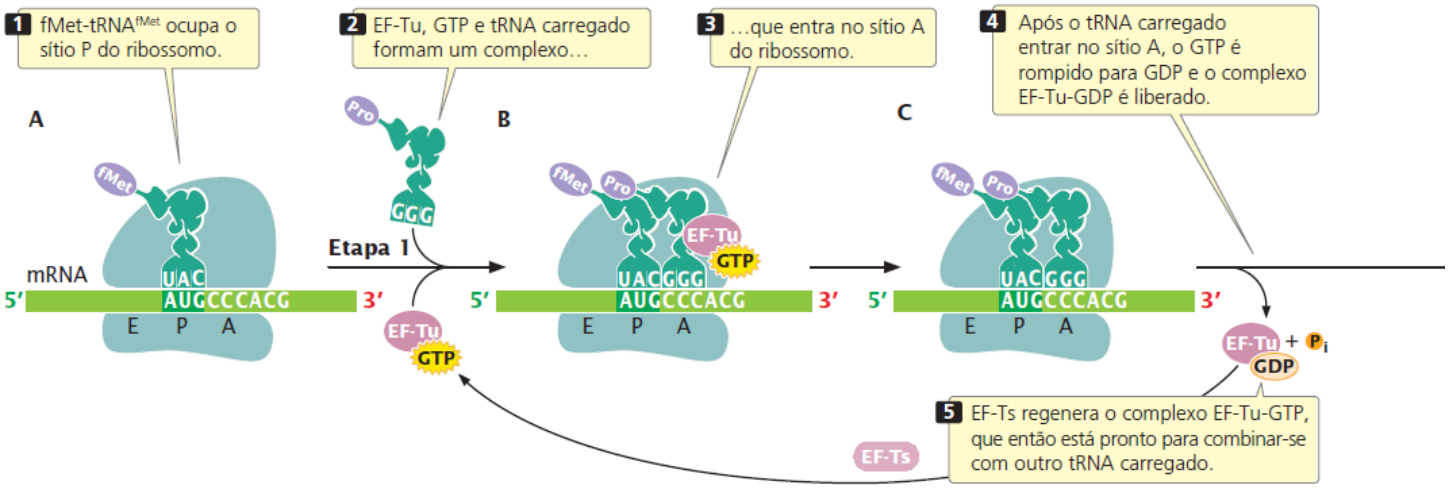


Figura 15.17 O alongamento da tradução inclui três etapas.

Após a translocação, o sítio A do ribossomo está vazio e pronto para receber o tRNA especificado para o próximo códon. O ciclo de alongamento (ver Figura 15.17 B a E) se repete: um tRNA carregado e seu aminoácido ocupam o sítio A, uma ligação peptídica é formada entre os aminoácidos nos sítios A e P, e o ribossomo se transloca para o próximo códon. Durante todo o ciclo, a cadeia de polipeptídios permanece presa ao tRNA no sítio P. Outra proteína, chamada fator P de alongamento translacional (EF-P), acentua a tradução das proteínas que têm cópias consecutivas do aminoácido prolina. Se EF-P estiver ausente, os ribossomos ficam parados durante a tradução dessas proteínas que têm poliprolina.

Os RNA mensageiros, embora sejam fitas simples, muitas vezes têm estruturas secundárias formadas pelo pareamento de bases complementares em partes diferentes do mRNA (ver Figura 13.1 B). À medida que o ribossomo se desloca ao longo do mRNA, essas estruturas secundárias são desenroladas pela atividade da helicase localizada na subunidade menor do ribossomo.

Recentemente, os pesquisadores desenvolveram métodos para seguir um único ribossomo à medida que ele traduz códons individuais de uma molécula de mRNA. Esses estudos revelaram que a tradução não ocorre de maneira contínua. Cada etapa de translocação requer tipicamente menos do que um décimo de segundo, mas às vezes existem pausas distintas, que amiúde duram alguns segundos, entre cada evento de translocação quando o ribossomo se desloca de um códon para outro. Assim, a tradução ocorre em várias translocações rápidas interrompidas por breves pausas. Além das pausas curtas entre as translocações, a tradução pode ser interrompida por pausas mais longas, de 1 a 2 minutos, que podem atuar na regulação do processo de tradução.

O alongamento nas células eucarióticas ocorre de forma semelhante. Os eucariotos têm pelo menos três fatores de alongamento, um dos quais atua no início e no término. Outro fator de alongamento usado nos eucariotos, chamado de fator 2 de alongamento eucariótico (eEF-2), é o alvo de uma toxina produzida pelas bactérias que causam difteria, uma doença que até pouco tempo era uma causa importante de morte de crianças. A toxina da difteria inibe o eEF-2, evitando a translocação do ribossomo ao longo do mRNA, e então a síntese de proteína para.

Conceitos

O alongamento consiste em três etapas: (1) um tRNA carregado entra no sítio A, (2) uma ligação peptídica é criada entre os aminoácidos nos sítios A e P, e (3) o ribossomo se transloca para o próximo códon. Este processo requer vários fatores de alongamento e GTP.

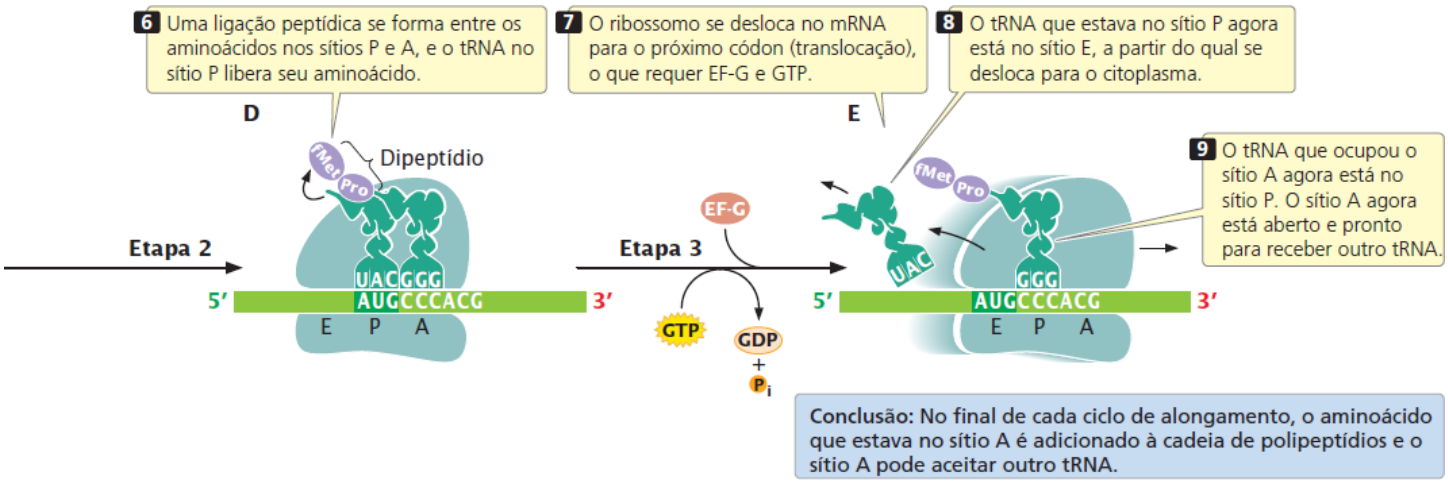
Checagem dos conceitos 8

No alongamento, a criação das ligações peptídicas entre os aminoácidos é catalisada por:

a. rRNA.
b. Proteína na subunidade menor.
c. Proteína na subunidade maior.
d. tRNA.

Terminação

A síntese de proteínas termina quando o ribossomo transloca para um códon de terminação. Como não existem tRNAs com complementaridade com anticódon no códon de terminação, nenhum tRNA entra no sítio A do ribossomo quando um códon de terminação é encontrado (**Figura 15.18 A**). Em vez disso, as proteínas chamadas de **fatores de liberação** se ligam ao ribossomo (**Figura 15.18 B**). *Escherichia coli* tem três fatores de liberação – RF-1, RF-2 e RF-3. O RF-1 se liga aos códons de terminação UUA e UAG, e o RF-2 se liga aos UGA e UAA. A ligação do fator de liberação RF-1 ou RF-2 ao sítio A do ribossomo promove a ruptura do tRNA no sítio P a partir da cadeia de polipeptídio e a liberação do polipeptídio. O RF-3 se liga ao ribossomo e forma um complexo com GTP. Esse complexo leva a uma mudança conformacional no ribossomo, liberando RF-1 ou RF-2 do sítio A e fazendo com que o tRNA no sítio P se desloque para o sítio E; no processo, GTP é hidrolisado a GDP. Fatores adicionais ajudam na liberação do tRNA do sítio P, a liberação do mRNA do ribossomo e a dissociação do ribossomo (**Figura 15.18 C**). É importante observar que o códon de terminação não está localizado na extremidade 3′ do mRNA; em vez disso, ele é seguido por vários nucleotídios que constituem a região 3′ não traduzida (UTR) do mRNA. A 3′ UTR frequentemente tem sequências que afetam a estabilidade do mRNA e influenciam a tradução (ver *Capítulo 17*).



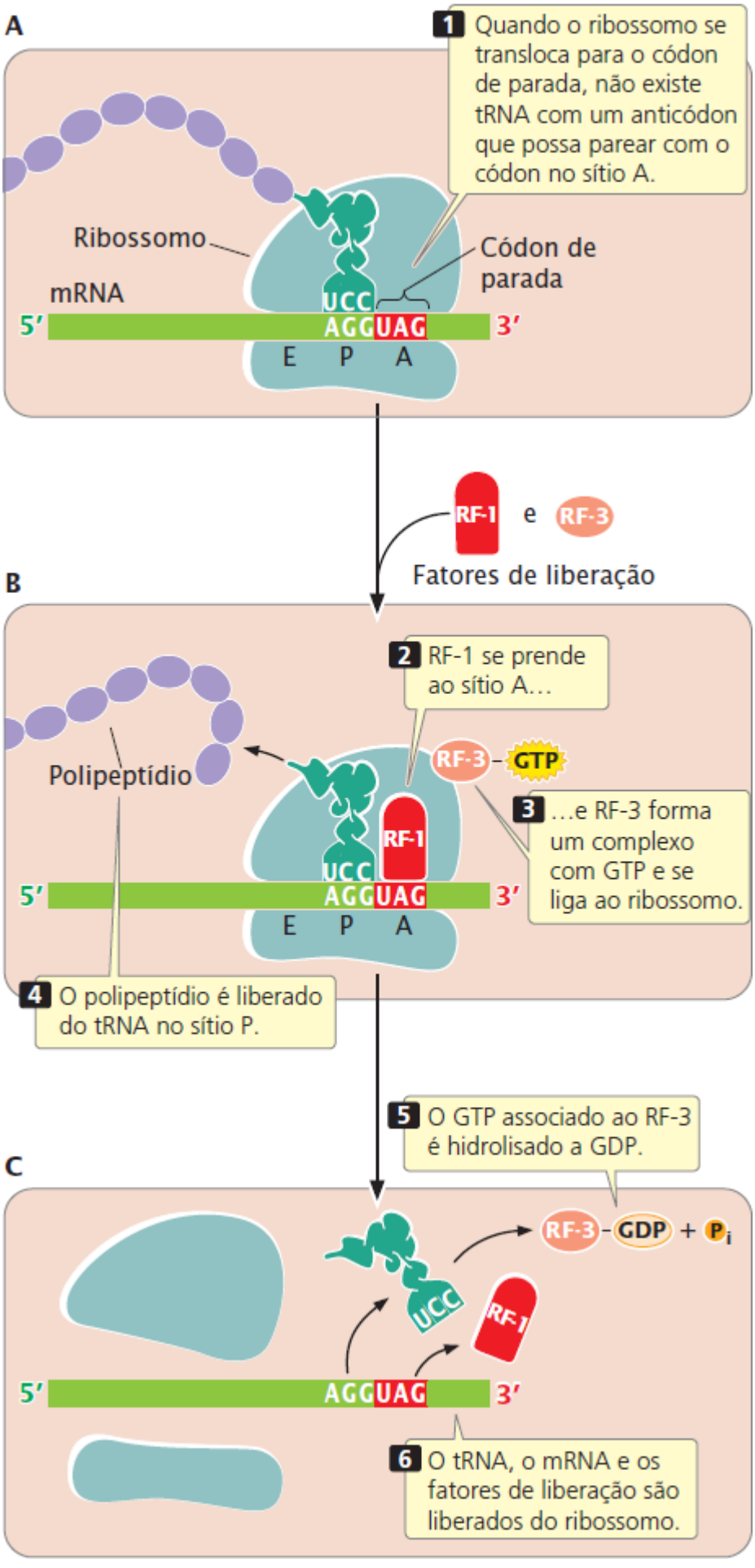


Figura 15.18 A translocação termina quando um códon de parada é encontrado. Como UAG é o códon de terminação nesta ilustração, o fator de liberação é RF-1.

Uma pesquisa recente mostra que alguns ribossomos bacterianos estão envolvidos em um tipo de revisão, semelhante à forma

que as DNA polimerases fazem a revisão durante a replicação. Após a translocação, o ribossomo verifica a interação entre o mRNA e o tRNA no sítio P. Se foi adicionado o tRNA errado, o alinhamento entre o mRNA e o tRNA estará incorreto, o que dispara o término prematuro da tradução. Evidências sugerem que uma função importante do RF-3 é ajudar a terminar a tradução quando for usado o tRNA errado.

A tradução nas células eucarióticas termina de forma semelhante, exceto pelo fato de que existem dois fatores de liberação: eRF-1, que reconhece todos os três códons de terminação, e eRF-2, que se liga ao GTP e estimula a liberação do polipeptídio do ribossomo. Não foi observada revisão da tradução nas células eucarióticas, como aquela estimulada pelo RF-3 nas bactérias.

Resolva os Problemas 30 e 34

Conceitos

A terminação ocorre quando o ribossomo alcança o códon de terminação. Os fatores de liberação se ligam ao códon de terminação, levando à liberação do polipeptídio do último tRNA, do tRNA do ribossomo e do mRNA do ribossomo.

A **Figura 15.19** resume o processo global de síntese de proteína, incluindo tRNA carregado, início, alongamento e término. O **Quadro 15.4** lista os componentes que participam desse processo.

Conceitos conectantes

Comparação da tradução bacteriana e da tradução eucariótica

Até o momento analisamos o processo de tradução nas células bacterianas e observamos algumas diferenças especiais que existem nas células eucarióticas. Vamos refletir sobre algumas semelhanças e diferenças importantes da síntese de proteínas nas células bacterianas e eucarióticas.

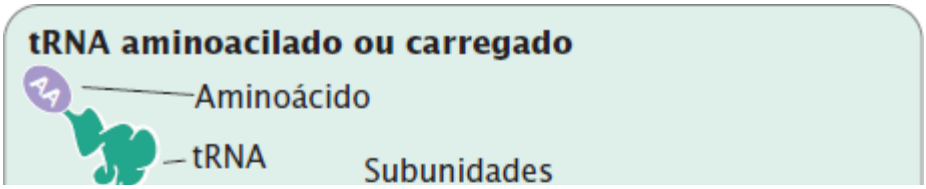
Primeiro, devemos enfatizar que o código genético da célula bacteriana e da célula eucariótica é praticamente idêntico; a única diferença está no aminoácido especificado pelo códon de iniciação. Nas células bacterianas, o códon AUG de início codifica um tipo modificado de metionina, *N*-formilmetionina, enquanto nas células eucarióticas o AUG de início codifica a metionina não formilada. Uma consequência do fato de as bactérias e os eucariotos usarem o mesmo código é que os genes eucarióticos podem ser traduzidos nos sistemas bacterianos e vice-versa; esse recurso possibilita o desenvolvimento da engenharia genética, como veremos no Capítulo 19 .

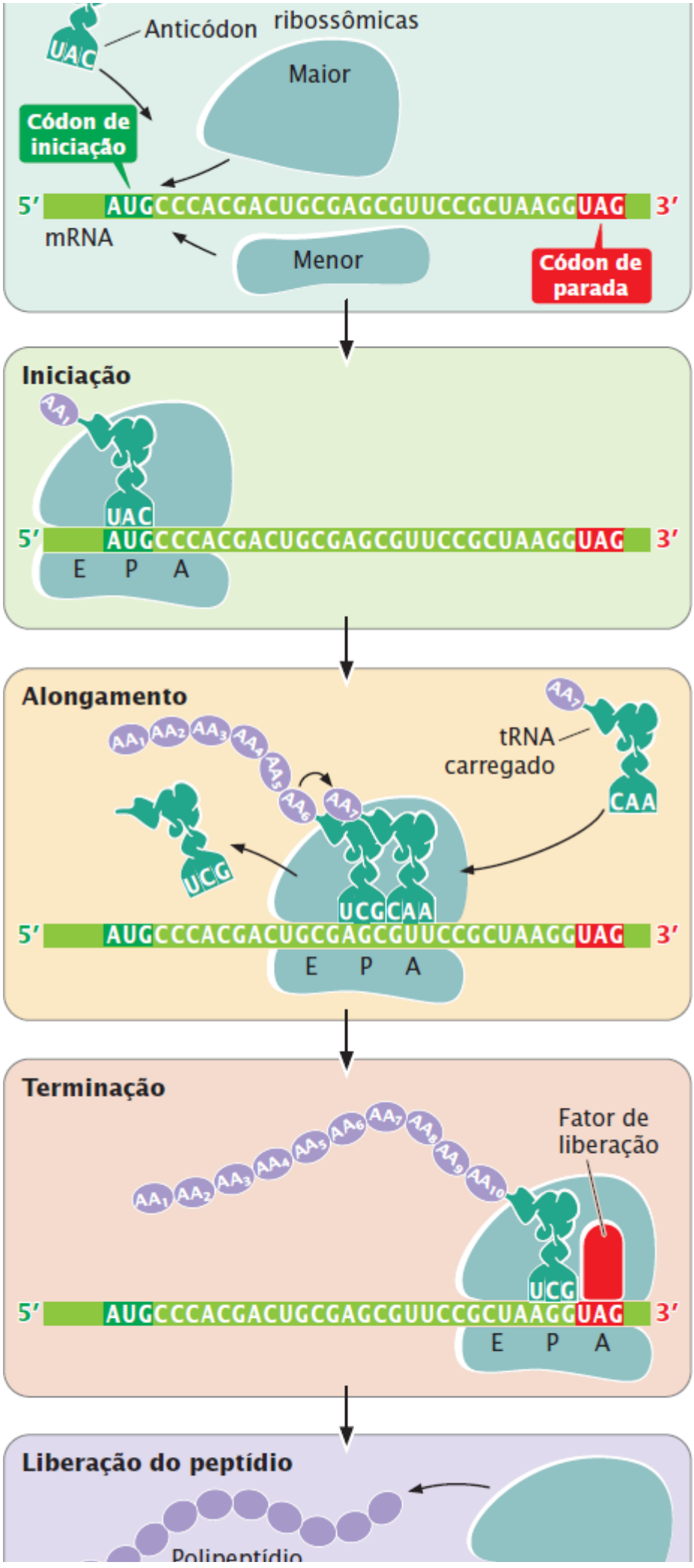
Outra diferença é que a transcrição e a tradução ocorrem simultaneamente nas células bacterianas, mas o envelope nuclear separa esses processos nas células eucarióticas. A separação física da transcrição e da tradução tem implicações importantes para o controle da expressão gênica, que vamos analisar no Capítulo 17 , e torna possível uma modificação abrangente nos mRNA eucarióticos, como foi discutido no Capítulo 14 .

Outra diferença é que o mRNA nas células bacterianas tem vida curta, durando apenas alguns minutos, mas o mRNA nas células eucarióticas pode durar horas ou dias. O cap 5′ e a cauda 3′ poli(A) encontrados nos mRNAs eucarióticos tornam possível sua estabilidade (ver *Capítulo 14*).

Tanto nas células bacterianas quanto nas eucarióticas, a aminoacil-tRNA sintetase liga os aminoácidos a seus tRNAs adequados e o processo químico é o mesmo. Existem diferenças importantes de tamanho e composição das subunidades ribossômicas bacterianas e eucarióticas. Por exemplo, a subunidade maior do ribossomo eucariótico tem três rRNAs, enquanto o ribossomo bacteriano tem apenas dois. Essas diferenças fazem com que os antibióticos e outras substâncias inibam a tradução bacteriana enquanto não têm efeito na tradução dos genes nucleares eucarióticos, como vamos discutir adiante neste capítulo.

Outras diferenças fundamentais estão no processo de iniciação. Nas células bacterianas, a subunidade menor do ribossomo se prende diretamente à região que circunda o códon de iniciação por meio de ponte de hidrogênio entre a sequência consenso Shine-Dalgarno na região 5′ não traduzida do mRNA e uma sequência na extremidade 3′ do 16S rRNA. Em vez disso, a subunidade menor de um ribossomo eucariótico se liga primeiro a proteínas fixas ao cap 5′ no mRNA e, então, migra pelo mRNA, “analisando” a sequência até encontrar o primeiro códon de iniciação AUG. Além disso, um número maior de fatores de iniciação atua na iniciação eucariótica do que na bacteriana.




Anticódon
UAC
ribossômicas
Maior
Códon de iniciação
5′ AUGCCCACGACUGCGAGCGUUCCGCUAAGGUAG 3′
mRNA
Menor
Códon de parada
Iniciação
AA1
UAC
5′ AUGCCCACGACUGCGAGCGUUCCGCUAAGGUAG 3′
E P A
Alongamento
AA1 AA2 AA3 AA4 AA5 AA6 AA7
tRNA carregado
CAA
UCG
UCG CAA
5′ AUGCCCACGACUGCGAGCGUUCCGCUAAGGUAG 3′
E P A
Terminação
AA1 AA2 AA3 AA4 AA5 AA6 AA7 AA8 AA9 AA10
Fator de liberação
UCG
5′ AUGCCCACGACUGCGAGCGUUCCGCUAAGGUAG 3′
E P A
Liberação do peptídio
Polipeptídio

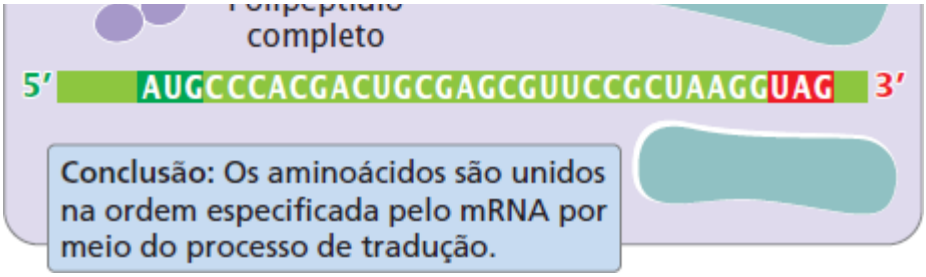


Figura 15.19 A tradução exige aminoacilação do tRNA, iniciação, alongamento e terminação. Neste processo, os aminoácidos são unidos na ordem especificada pelo mRNA para criar uma cadeia de polipeptídios. Vários fatores de iniciação, alongamento e liberação participam do processo, e a energia é fornecida por ATP e GTP.

Alongamento e término são semelhantes nas células bacterianas e eucarióticas, embora sejam usados diferentes fatores de alongamento e término. Em ambos os tipos de organismos, os mRNAs são traduzidos múltiplas vezes e são simultaneamente presos a vários ribossomos, formando polirribossomos, como discutido na Seção 15.4.

Pouco se sabe sobre o processo de tradução em archaea, mas parece que eles têm uma mistura das características eubacterianas e eucarióticas. Como as archaea não têm membrana nuclear, a transcrição e a tradução ocorrem simultaneamente, como nas células eubacterianas. Archaea usam metionina não formilada como o aminoácido iniciador, uma característica da tradução eucariótica. Alguns dos fatores de iniciação e liberação em archaea são semelhantes aos encontrados nas eubactérias, enquanto outros são semelhantes aos encontrados nos eucariotos. Finalmente, alguns dos antibióticos que inibem a tradução nas eubactérias não têm efeito na tradução em archaea, fornecendo evidência das diferenças fundamentais entre eubactérias e archaea.

15.4 Propriedades adicionais do RNA e dos ribossomos influenciam a síntese de proteínas

Agora que vimos em detalhes o processo de tradução, vamos examinar alguns aspectos adicionais da síntese de proteínas e seu maquinário.

Estrutura tridimensional do ribossomo

O papel central do ribossomo na síntese de proteína foi identificado nos anos de 1950, e numerosos aspectos da sua estrutura foram estudados desde então. Contudo, muitos aspectos permaneceram um mistério até que as reconstruções tridimensionais detalhadas fossem completadas. Em 2009, Venkatraman Ramakrishnan, Thomas Steitz e Ada Yonath receberam o Prêmio Nobel de Química por sua pesquisa na estrutura molecular do ribossomo.

A **Figura 15.20 A** mostra um modelo de ribossomo de *E. coli* com baixa resolução. Uma imagem de alta resolução do ribossomo bacteriano como determinado pela cristalografia por raios X está representada no modelo descrito na **Figura 15.20 B**. O mRNA é ligado à subunidade menor do ribossomo e os tRNA estão localizados nos sítios A, P e E (**Figura 15.20 C**) que fazem a ponte entre as subunidades menor e maior (ver Figura 15.20 A). Os fatores de iniciação 1 e 3 se ligam aos sítios no lado de fora da subunidade menor do ribossomo. EF-Tu, EF-G e outros fatores complexados com GTP interagem com o *centro de ligação do fator*. As imagens de cristalografia de alta resolução fornecem informações indicando que o *centro decodificador* reside na subunidade menor do ribossomo (o centro decodificador não pode ser observado na Figura 15.20 B). Esse centro detecta o encaixe entre o códon no mRNA e o anticódon no tRNA carregado que chega. Apenas os tRNAs com o anticódon correto estão firmemente ligados ao ribossomo. As análises estruturais também indicam que a subunidade maior do ribossomo tem o *centro da peptidil transferase*, onde ocorre a formação da ligação peptídica. Não existem proteínas ribossômicas no centro da peptidil transferase; a formação da ligação peptídica é feita pela molécula de RNA. Essas análises também revelam que um túnel conecta o sítio da formação da ligação peptídica com a parte posterior do ribossomo; a cadeia polipeptídica crescente passa por esse túnel para o lado externo do ribossomo. O túnel consegue acomodar cerca de 35 aminoácidos da cadeia polipeptídica crescente.

Quadro 15.4 Componentes necessários para a síntese proteica nas células bacterianas.

Estágio	Componente	Função
tRNA carregado	Aminoácidos	Blocos construtores de proteínas

	tRNAs	Entregam aminoácidos aos ribossomos
	Aminoacil-tRNA sintetases	Ligam aminoácidos aos tRNAs
	ATP	Fornece energia para ligar o aminoácido ao tRNA
Iniciação	mRNA	Carrega as instruções de codificação
	fMet-tRNA^{fMet}	Fornece o primeiro aminoácido no peptídio
	Subunidade ribossômica 30S	Fixa-se ao mRNA
	Subunidade ribossômica 50S	Estabiliza os tRNAs e aminoácidos
	Fator de iniciação 1	Potencializa a dissociação de subunidades maior e menor do ribossomo
	Fator de iniciação 2	Liga-se ao GTP; entrega fMet-tRNA^{fMet} ao códon de iniciação
	Fator de iniciação 3	Liga-se à subunidade 30S e evita a associação com a subunidade 50S
Alongamento	Complexo de iniciação 70S	Ribossomo funcional com os sítios A, P e E onde ocorre a síntese de proteínas
	tRNAs carregados	Trazem os aminoácidos para o ribossomo e ajudam a montá-los na ordem especificada pelo mRNA
	Fator de alongamento Tu	Liga-se ao GTP e tRNA carregado; entrega o tRNA carregado ao sítio A
	Fator de alongamento Ts	Gera o fator Tu de alongamento ativo
	Fator de alongamento G	Estimula o movimento do ribossomo para o próximo códon
	GTP	Fornece energia
	Centro da peptidil transferase	Cria a ligação peptídica entre aminoácidos nos sítios A e P
Terminação	Fatores de liberação 1, 2 e 3	Ligam-se ao ribossomo quando o códon de parada é alcançado e terminam a tradução

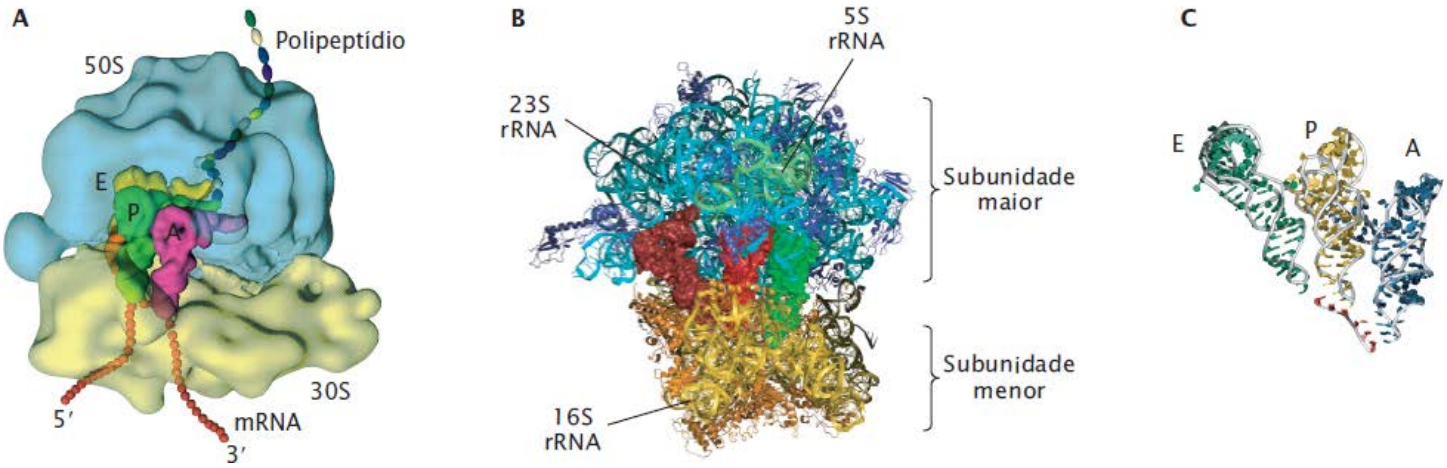


Figura 15.20 A estrutura do ribossomo. A. Modelo de baixa resolução do ribossomo, mostrando os sítios A, P e E onde estão os tRNAs, o mRNA e a cadeia de polipeptídios crescente. **B.** Modelo de alta resolução do ribossomo. O tRNA no sítio E é marrom-

avermelhado, o tRNA no sítio P é vermelho e o tRNA no sítio A é verde. O 16S rRNA é apresentado em amarelo, o 23S rRNA em ciano e o 5S rRNA em verde. As proteínas da subunidade ribossômica menor são apresentadas em laranja, enquanto as proteínas da subunidade ribossômica maior são apresentadas em púrpura. **C.** Posições dos tRNAs nos sítios E, P e A do ribossomo apresentado na parte B. (*Parte B*: MRC Lab of Molecular Biology, Wellcome Images.)

Polirribossomos

Nas células procarióticas e eucarióticas, as moléculas de mRNA são traduzidas simultaneamente por múltiplos ribossomos (**Figura 15.21**). A estrutura resultante – um mRNA com vários ribossomos presos – é chamada de **polirribossomo** (ou apenas polissomo). Cada ribossomo se prende sucessivamente ao sítio de ligação do ribossomo na extremidade 5′ do mRNA e se move no sentido da extremidade 3′; o polipeptídio associado a cada ribossomo se torna progressivamente maior à medida que o ribossomo se move ao longo do mRNA. Nas células procarióticas, a transcrição e a tradução são simultâneas; múltiplos ribossomos podem ser ligados à extremidade 5′ do mRNA enquanto a transcrição ainda está ocorrendo na extremidade 3′, como apresentado na Figura 15.21. Nos eucariotos, a transcrição e a tradução são separadas no tempo e espaço, com a transcrição ocorrendo no núcleo e a tradução ocorrendo no citoplasma.

Conceitos

Nas células procarióticas e eucarióticas, múltiplos ribossomos podem ser ligados a um único mRNA, gerando uma estrutura chamada de polirribossomo.

Checagem dos conceitos 9

Em um polirribossomo, os polipeptídios associados a quais ribossomos serão os maiores?

a. Os que estão na extremidade 5′ do mRNA.
b. Os que estão na extremidade 3′ do mRNA.
c. Os que estão no meio do mRNA.
d. Todos os polipeptídios terão o mesmo comprimento.

Vigilância do RNA mensageiro

A transferência exata da informação genética de uma geração para a próxima e do genótipo para o fenótipo é crítica para o desenvolvimento e o funcionamento adequados de um organismo. Consequentemente, as células desenvolveram vários mecanismos de controle de qualidade para garantir a precisão da transferência das informações. A síntese de proteínas não é exceção: vários mecanismos, coletivamente chamados de **vigilância do mRNA**, existem para detectar e tratar dos erros nos mRNAs que podem criar problemas no curso da tradução. Esses mecanismos evitam que a célula perca recursos traduzindo mRNAs aberrantes e previnem a produção de proteínas truncadas, que podem ser tóxicas para a célula.

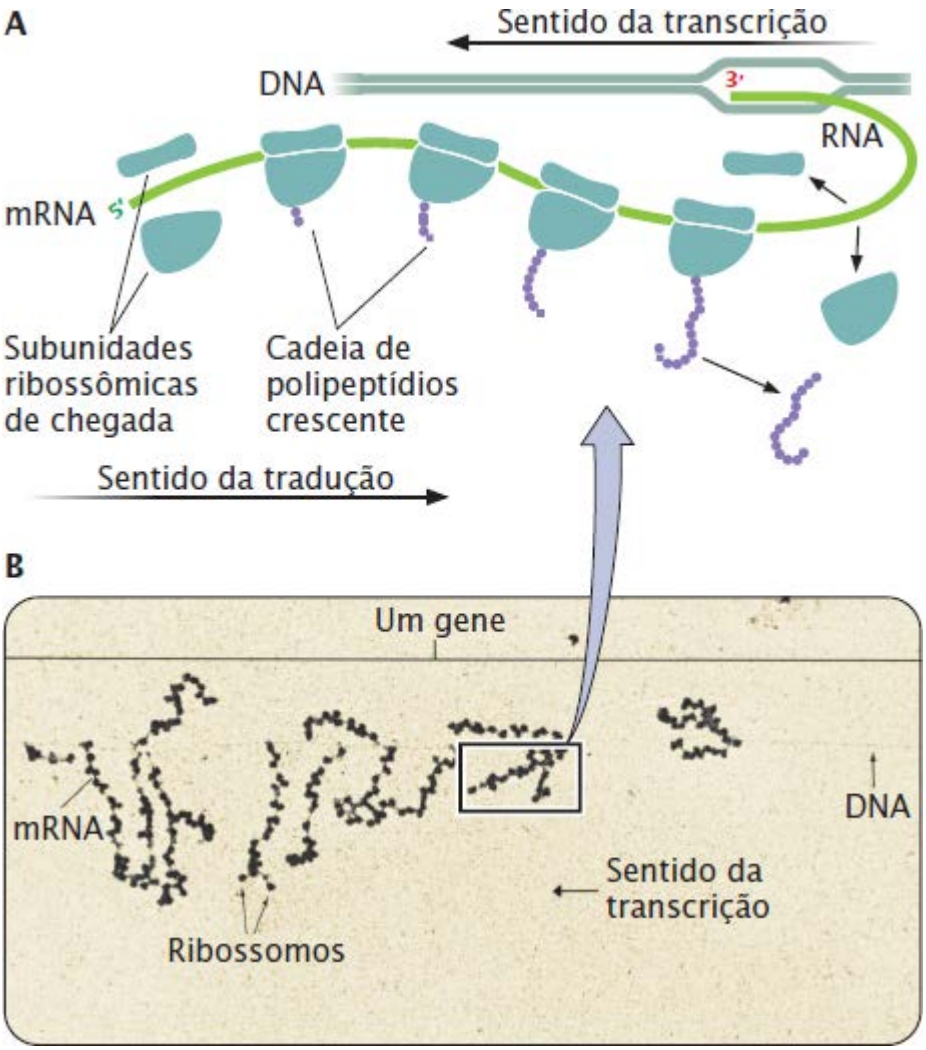


Figura 15.21 Uma molécula de mRNA pode ser traduzida simultaneamente por vários ribossomos. A. Quatro ribossomos são traduzidos em uma molécula de mRNA; os ribossomos se deslocam da extremidade 5′ para a extremidade 3′ do mRNA. **B.** Nesta eletromicrografia, o longo filamento horizontal é o DNA, as esferas escuras são os polirribossomos e os filamentos finos presos aos ribossomos são os mRNAs. A transcrição do DNA segue da direita para esquerda; os mRNA da direita são mais curtos que os da esquerda. Cada mRNA está sendo traduzido por múltiplos ribossomos. (*Parte B*: O. L. Miller, Jr., and B.A. Hamaklo, and C.A. Thomas, Jr. Science 169(1970):392. Reimpressa com permissão de AAAS.)

Decaimento do mRNA mediado por mutação sem sentido. Uma mutação comum é aquela na qual um códon que especifica um aminoácido é alterado para se tornar um códon de terminação (chamada de mutação sem sentido ou *nonsense*; ver *Capítulo 18*). Uma mutação sem sentido não afeta a transcrição, mas a tradução termina de forma prematura quando o códon de terminação é encontrado. A proteína resultante é truncada e em muitos casos, não funcional. Um modo comum de surgimento das mutações sem sentido nas células eucarióticas é quando um ou mais éxons são saltados ou incorretamente recompostos. O *splicing* incorreto leva à deleção ou à adição de nucleotídios no mRNA, que altera o quadro de leitura e introduz códons de terminação prematuros.

Para evitar a síntese de proteínas aberrantes resultantes de mutação sem sentido, as células eucarióticas desenvolveram um mecanismo de **decaimento de mRNA mediado por terminação sem sentido (*nonsense*)** (NMD, do inglês, *nonsense-mediated mRNA decay*), que leva à rápida eliminação do mRNA contendo códons de terminação prematuros. O mecanismo responsável pelo NMD é pouco compreendido. Nos mamíferos, parece que vincula proteínas que se ligam a junções éxon-éxon. Essas proteínas de junção de éxon podem interagir com enzimas que degradam o mRNA. Uma possibilidade é que o primeiro ribossomo a traduzir o mRNA (no ciclo pioneiro de tradução) remove as proteínas de junção de éxon, protegendo, assim, o mRNA da degradação. Entretanto, quando o ribossomo encontra um códon de terminação prematuro, ele não cruza o mRNA inteiro, e algumas proteínas de junção de éxon não são removidas, resultando em decaimento de mRNA mediado por mutação sem sentido. Resolva o Problema 38

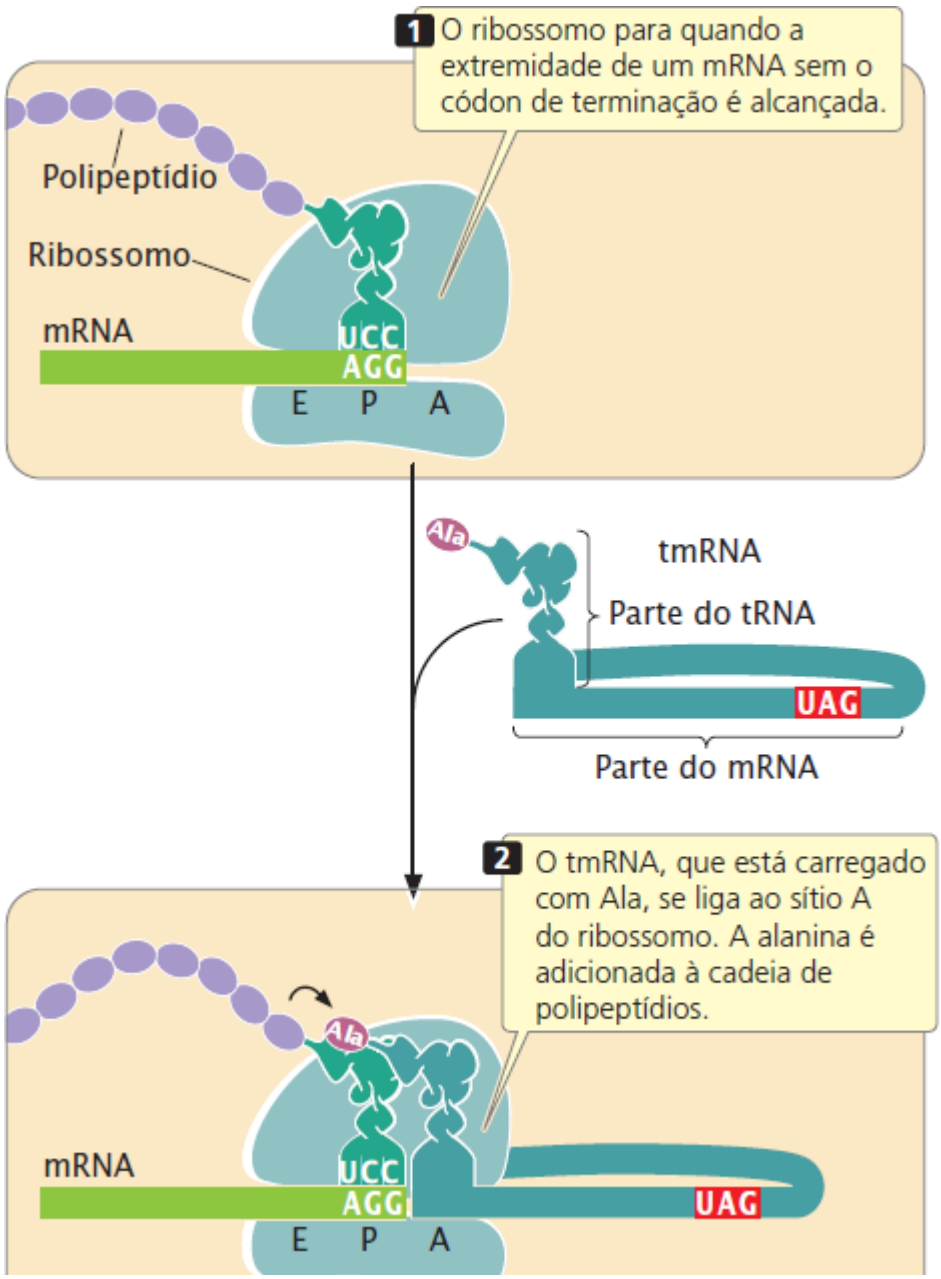
Ribossomos parados e mRNAs nonstop. Um problema que surge na tradução é quando um ribossomo para no mRNA antes de terminar a tradução. Essa situação pode surgir quando uma mutação no DNA altera um códon de terminação em um códon que especifica um aminoácido. Ela também pode surgir quando a transcrição termina de forma prematura, produzindo um mRNA truncado sem um códon de terminação. Nesses casos, o ribossomo alcança a extremidade do mRNA sem encontrar

um códon de terminação e para, ainda preso ao mRNA. A fixação ao mRNA impede que o ribossomo seja reciclado para ser usado em outros mRNAs e, se tais ocorrências forem frequentes, o resultado é a falta de ribossomos, o que diminui os níveis gerais de síntese de proteína.

As bactérias desenvolveram um "caminhão de reboque" molecular, chamado de **RNA transportador-mensageiro** (tmRNA), que remove os ribossomos parados. Essa molécula de RNA tem propriedades de tRNA e mRNA; seu componente tRNA está carregado com o aminoácido alanina. Quando um ribossomo fica preso em um mRNA, EF-Tu entrega o tmRNA para o sítio A do ribossomo, onde o tmRNA age como um tRNA substituto (**Figura 15.22**). Uma ligação peptídica é criada entre o aminoácido no sítio P do ribossomo e da alanina (ligada ao tmRNA) agora no sítio A, transferindo a cadeia de polipeptídio para o tmRNA e liberando o tRNA no sítio P.

O ribossomo então reinicia a tradução, trocando o mRNA original aberrante para a parte mRNA do tmRNA. A tradução adiciona 10 aminoácidos codificados pelo tmRNA, e então um códon de terminação é alcançado na extremidade 3′ do tmRNA, que termina a tradução e libera o ribossomo. Os aminoácidos adicionados são um marcador da cadeia de polipeptídio incompleta a ser degradado. Algumas evidências sugerem que o tmRNA também direciona o mRNA aberrante para degradação. Não está claro como os ribossomos paralisados são reconhecidos pelo tmRNA, mas esse método é eficiente na reciclagem de ribossomos paralisados e eliminar proteínas anormais que surgem a partir da transcrição truncada.

Os eucariotos desenvolveram um mecanismo diferente para lidar com o mRNA que não tem códons de terminação. Em vez de reiniciar o ribossomo parado e degradar a proteína anormal resultante, as células eucarióticas usam um mecanismo chamado de **decaimento de mRNA *nonstop* (sem fim)**, que resulta na rápida degradação do mRNA anormal. Nesse mecanismo, o sítio A livre de códon do ribossomo paralisado é reconhecido por uma proteína especial que se liga a este sítio e recruta outras proteínas, que então degradam o mRNA a partir da sua extremidade 3′.



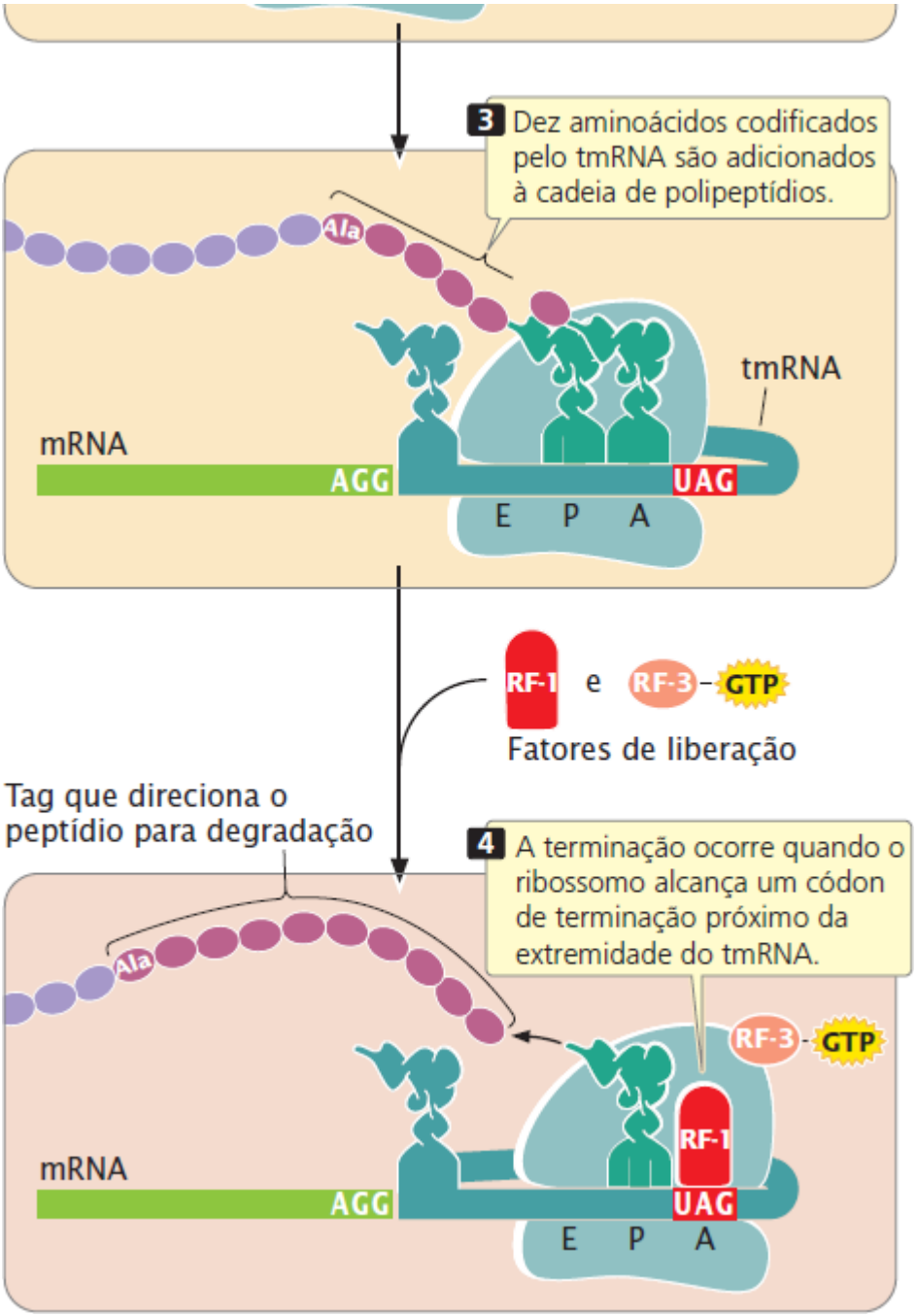


Figura 15.22 O tmRNA nas bactérias possibilita que os ribossomos parados reiniciem a tradução.

Decaimento *no-go*. Outro sistema de vigilância do mRNA encontrado nos eucariotos é o **decaimento *no-go*** (NGD), que ajuda a remover ribossomos parados oriundos de estruturas secundárias no mRNA, dano químico ao mRNA, códons de parada prematura e defeitos nos ribossomos. Uma série de proteínas leva a terminação, reciclagem dos ribossomos e degradação do mRNA.

Conceitos

As células têm mecanismos de vigilância do mRNA para detectar e eliminar moléculas de mRNA com erros que comprometem a tradução.

Dobradura e modificações pós-tradução das proteínas

As funções de muitas proteínas dependem em muito da dobradura correta da cadeia de polipeptídios; algumas proteínas se dobram espontaneamente em seus formatos corretos, mas, para outras, a dobra correta pode requerer a princípio a participação de outras moléculas, chamadas de **chaperonas moleculares**. Algumas chaperonas moleculares estão associadas ao ribossomo e dobram cadeias de polipeptídio recém-sintetizadas à medida que elas surgem do túnel do ribossomo, no qual a dobra da proteína ocorre à medida que a tradução avança.

Muitas proteínas devem ser modificadas após a tradução para se tornarem ativas. As proteínas nas células procarióticas e eucarióticas sofrem alterações após a tradução, que são chamadas de modificações pós-tradução. São possíveis vários tipos diferentes de modificações. Algumas proteínas são sintetizadas como proteínas precursoras maiores e devem ser rompidas e aparadas por enzimas antes de se tornarem funcionais. Como já foi mencionado, o grupo formil ou o resíduo inteiro de metionina podem ser removidos da extremidade amino de uma proteína. Algumas proteínas exigem a fixação de carboidratos para ativação. Os aminoácidos dentro de uma proteína podem ser modificados: são adicionados fosfatos, grupos carboxila e grupos metila em alguns aminoácidos. Nas células eucarióticas, a extremidade amino de uma proteína é acetilada após a tradução.

Uma modificação pós-tradução comum nos eucariotos é a fixação de uma proteína chamada ubiquitina, que direciona a proteína para degradação. Outra modificação de algumas proteínas é a remoção de 15 a 30 aminoácidos, chamada de **sequência de sinalização** na extremidade amino da proteína. A sequência de sinalização ajuda a direcionar uma proteína para um local específico dentro da célula, após o que a sequência é removida por enzimas especiais.

Conceitos

Muitas proteínas sofrem modificações pós-tradução depois de sua síntese.

Tradução e antibióticos

Os antibióticos são fármacos usados para exterminar microrganismos. Para criar um antibiótico efetivo, não um veneno qualquer, o truque é matar as bactérias sem prejudicar o paciente.

A tradução é frequentemente o alvo dos antibióticos porque é essencial para todos os organismos vivos e apresenta diferenças significativas entre as células bacterianas e eucarióticas. Vários antibióticos se ligam seletivamente a ribossomos bacterianos e inibem diversas etapas na tradução, mas eles não afetam os ribossomos eucarióticos. As tetraciclinas, por exemplo, são uma classe de antibióticos que se ligam ao sítio A do ribossomo bacteriano e bloqueiam a entrada de tRNA carregados, visto que eles não têm efeito nos ribossomos eucarióticos. A neomicina se liga ao ribossomo próximo do sítio A e induz a erros de tradução, provavelmente ao provocar erros na ligação de tRNAs carregados ao sítio A. O cloranfenicol se liga à subunidade maior do ribossomo e bloqueia a formação da ligação peptídica. A estreptomicina se liga à subunidade menor do ribossomo e inibe o início, e a eritromicina bloqueia a translocação. Embora o cloranfenicol e a estreptomicina sejam potentes inibidores da tradução nas bactérias, eles não inibem a tradução em archaea.

A estrutura tridimensional da puromicina lembra a extremidade 3′ do tRNA carregado, tornando possível que a puromicina entre no sítio A de um ribossomo de forma eficiente e iniba a entrada de tRNAs. Uma ligação peptídica pode se formar entre a molécula de puromicina no sítio A e um aminoácido no tRNA no sítio P do ribossomo, mas a puromicina não pode se ligar ao sítio P e a translocação não ocorre, bloqueando o alongamento da proteína. Como a estrutura do tRNA é semelhante em todos os organismos, a puromicina inibe a tradução nas células bacterianas e eucarióticas; consequentemente, ela extermina as células eucarióticas junto com as bactérias e, às vezes, é usada para destruir as células cancerosas.

Muitos antibióticos bloqueiam etapas específicas na tradução, e diferentes antibióticos bloqueiam diferentes etapas na síntese proteica, como o início ou o alongamento. Por causa dessa especificidade, os antibióticos são usados para estudar o processo de síntese de proteínas.

Resumo dos conceitos

- George Beadle e Edward Tatum desenvolveram a hipótese de um gene, uma enzima, que propunha que cada gene especifica uma enzima; esta hipótese foi posteriormente modificada para se tornar a hipótese de um gene, um polipeptídio
- As proteínas são compostas por 20 aminoácidos diferentes. Os aminoácidos em uma proteína são unidos por meio de ligações peptídicas. As cadeias de aminoácidos se dobram e se associam para formar as estruturas secundárias, terciárias e quaternárias das proteínas
- Solucionar o código genético exigiu diferentes tipos de abordagens, incluindo o uso de mRNAs sintéticos com sequências aleatórias e mRNAs curtos que se ligavam a tRNAs carregados
- O código genético é um código de trincas: três nucleotídios especificam um único aminoácido. Ele também é degenerado (o que significa que mais de um códon pode especificar um aminoácido), não sobreposto e (quase) universal
- Diferentes tRNAs (tRNAs isoaceptores) podem aceitar o mesmo aminoácido. Diferentes códons podem parear com o mesmo anticódon pelo *wobble* (oscilação), que pode existir na terceira posição do códon e torna possível o pareamento não padrão de bases nessa posição
- O quadro de leitura é determinado pelo códon de iniciação. A extremidade da seção codificadora de proteína de um mRNA é

marcada por um de três códons de terminação
- A síntese proteica inclui quatro etapas: (1) a ligação de aminoácidos nos tRNAs adequados, (2) início, (3) alongamento e (4) terminação
- A ligação de um aminoácido a um tRNA requer uma aminoacil-tRNA sintetase específica e ATP. O aminoácido é fixo por sua extremidade carboxila à extremidade 3′ do tRNA
- No início da tradução bacteriana, a subunidade menor do ribossomo se fixa ao mRNA e é posicionada sobre o códon de iniciação. Ela é unida pelo primeiro tRNA e seu aminoácido associado (*N*-formilmetionina nas células bacterianas) e, posteriormente, pela subunidade maior do ribossomo. O início requer vários fatores de iniciação e GTP
- No alongamento, um tRNA carregado entra no sítio A de um ribossomo, uma ligação peptídica é criada entre os aminoácidos nos sítios A e P e o ribossomo se move (transloca) no mRNA para o próximo códon. Esse processo requer vários fatores de alongamento e GTP
- A tradução é terminada quando o ribossomo encontra um dos três códons de terminação. São necessários os fatores de liberação e GTP para causar a terminação
- Cada mRNA pode ser simultaneamente traduzido por vários ribossomos, produzindo uma estrutura chamada de polirribossomo
- As células têm mecanismos de vigilância do RNA que eliminam mRNAs com erros que podem comprometer a tradução
- Os antibióticos atuam ao interferir na tradução, porque muitos aspectos da tradução das bactérias são diferentes da tradução nos eucariotos
- Muitas proteínas sofrem modificação pós-tradução.

Termos importantes

Aminoácido
Aminoacil-tRNA sintetase
Chaperona molecular
Código genético degenerado
Código genético não sobreposto
Código genético universal
Códon de iniciação
Códon de terminação (parada ou *nonsense*)
Códon com sentido (senso)
Códons sinônimos
Complexo de iniciação 30S
Complexo de iniciação 70S
Complexo de ligação ao cap (CBC)
Decaimento de RNA *nonstop* (sem fim)
Decaimento do mRNA mediado pela terminação sem sentido (*nonsense*) (NMD)
Decaimento *no-go*
Fator G de alongamento (EF-G)
Fator Ts de alongamento (EF-Ts)
Fator Tu de alongamento (EF-Tu)
Fator de iniciação (IF-1, IF-2, IF-3)
Fator de liberação (RF-1, RF-2, RF-3)
Hipótese de um gene, uma enzima
Hipótese de um gene, um polipeptídio
Ligação peptídica
Oscilação (*wobble*)
Polipeptídio
Polirribossomo
Quadro de leitura
RNA transportador-mensageiro (tmRNA)
Sequência de sinalização
Sítio aminoacil (A)
Sítio de saída (E)
Sítio de peptidil (P)
Translocação

tRNA carregado
tRNAs isoaceptores
Vigilância do mRNA

Respostas da Checagem dos conceitos

1. B → A → C.
2. A sequência de aminoácidos (estrutura primária) de uma proteína.
3. b.
4. d.
5. O códon de iniciação nas bactérias codifica *N*-formilmetionina; nos eucariotos, ele codifica a metionina. Os códons de terminação não especificam aminoácidos.
6. c.
7. A sequência de Shine-Dalgarno.
8. a.
9. b.

Problemas desenvolvidos

Problema 1

Vários mutantes auxotróficos foram isolados em *Neurospora*. O exame dos fungos com essas mutações revelou que eles cresceram no meio mínimo ao qual foram adicionados vários compostos (A, B, C, D); as respostas de crescimento a cada um dos quatro compostos estão no quadro a seguir. Indique a ordem dos compostos A, B, C e D em uma via bioquímica. Esboce uma via bioquímica que inclua esses quatro compostos e indique que etapa na via é afetada por cada uma das mutações.

	Composto			
Número da mutação	**A**	**B**	**C**	**D**
134	+	+	−	+
276	+	+	+	+
987	−	−	−	+
773	+	+	+	+
772	−	−	−	+
146	+	+	−	+
333	+	+	−	+
123	−	+	−	+

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
A ordem dos compostos em uma via bioquímica; para cada mutação, qual etapa na via é afetada pela mutação.

Quais informações são fornecidas para solucionar o problema?
Para cada mutante, se ele cresceu no meio mínimo ao qual os compostos A, B, C e D foram adicionados.

Para a solução deste problema, revise:
Hipótese de um gene, uma enzima, na Seção 15.1.

Etapas para a solução

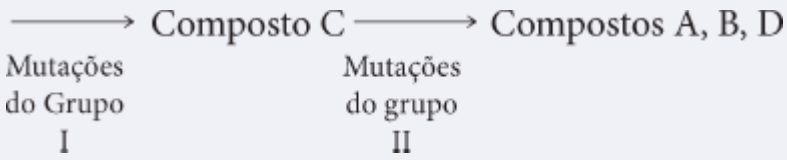
Mutação		Composto			
Grupo	**Número**	**A**	**B**	**C**	**D**
I	276	+	+	+	+
	773	+	+	+	+
II	134	+	+	−	+
	146	+	+	−	+
	333	+	+	−	+
III	123	−	+	−	+
IV	987	−	−	−	+
	772	−	−	−	+

Dica: Agrupe as mutações a partir de quais compostos tornam possível o crescimento.

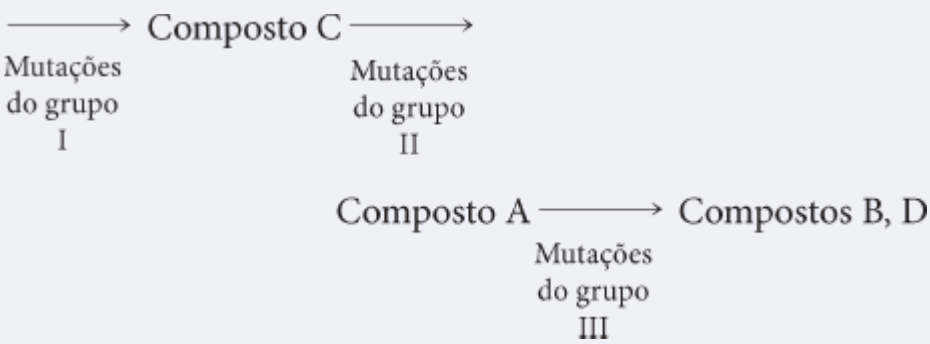
Os mutantes no grupo I crescerão se os compostos A, B, C ou D forem adicionados ao meio; então, essas mutações influenciam obrigatoriamente uma etapa antes da produção de todos os compostos:

Dica: Se um composto for adicionado após o bloqueio, ele possibilitará que o mutante cresça; se um composto for adicionado antes do bloqueio, ele não terá efeito.

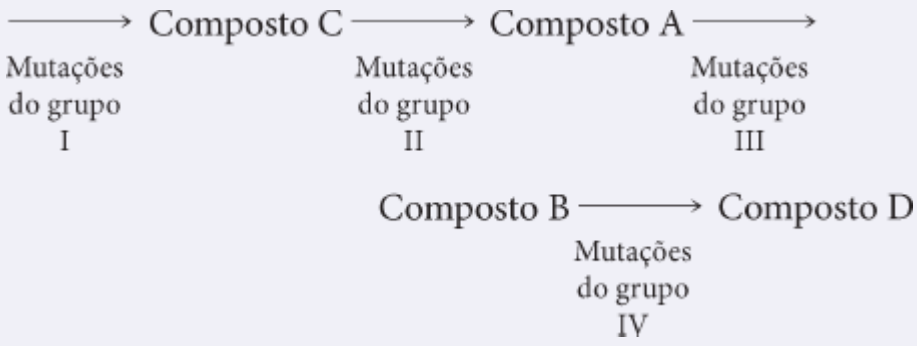
Os mutantes do grupo II crescerão se os compostos A, B ou D forem adicionados, mas não se o composto C for adicionado. Assim, o composto C vem antes de A, B e D; e as mutações do grupo II afetam a conversão do composto C em um dos outros compostos:



Os mutantes do grupo III crescerão se os compostos B ou D forem adicionados, mas não se os compostos A ou C forem adicionados. Dessa forma, as mutações do grupo III afetam as etapas que seguem a produção de A e C; já determinamos que o composto C precede A na via, e então A tem de ser o próximo composto na via:



Finalmente, os mutantes no grupo IV crescerão se o composto D for adicionado, mas não se os compostos A, B ou C forem adicionados. Assim, o composto D é o quarto composto na via, e as mutações no grupo IV bloqueiam a conversão de B em D:



Problema 2

Uma fita molde no DNA bacteriano tem a seguinte sequência:

5′-AGGTTTAACGTGCAT-3′

Quais aminoácidos são codificados por esta sequência?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
A lista de aminoácidos codificada pela sequência fornecida.

Quais informações são fornecidas para solucionar o problema?

- A sequência de DNA da fita molde
- As extremidades 5′ e 3′ da sequência molde
- Os aminoácidos codificados por diferentes códons (Figura 15.8).

Para a solução deste problema, revise:
A degeneração do código, na Seção 15.2.

Etapas para a solução

Para responder a esta pergunta, primeiro é preciso trabalhar na sequência mRNA que será transferida a partir dessa sequência de DNA:

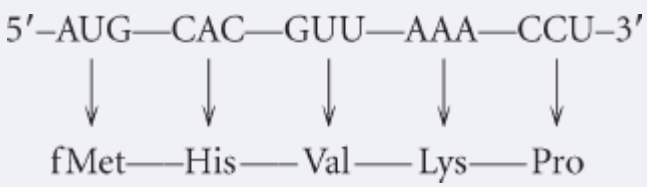
Lembrete: O mRNA é antiparalelo e complementar à fita molde de DNA.

Fita molde de DNA:	5′-AGGTTTAACGTGCAT-3′
mRNA copiado a partir do DNA:	3′-UCCAAAUUGCACGUA-5′

Um mRNA é traduzido no sentido 5′ → 3′; então, é útil se girarmos a molécula de RNA ao redor com a extremidade 5′ para a esquerda:

mRNA copiado do DNA:	5′-AUGCACGUUAAACCU-3′

Os códons são grupos de três nucleotídios que são lidos com sucesso após o primeiro códon AUG; considerando a Figura 15.8, podemos determinar que os aminoácidos são



Problema 3

As trincas a seguir constituem anticódons encontrados em vários tRNAs. Nomeie o aminoácido carregado por cada um desses tRNAs.

a. 5′-UUU-3′.
b. 5′-GAC-3′.
c. 5′-UUG-3′.
d. 5′-CAG-3′.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
O aminoácido carregado por cada tRNA.

Quais informações são fornecidas para solucionar o problema?

- A sequência do anticódon de cada tRNA
- Os aminoácidos codificados por diferentes códons (Figura 15.8).

Para a solução deste problema, revise:
A degeneração do código, na Seção 15.2.

Etapas para a solução

Para solucionar este problema, primeiro determinamos os códons com os quais esses anticódons pareiam, e então procuramos o aminoácido especificado pelo códon na Figura 15.8 . Para parte *a*, o anticódon é 5′-UUU-3′. De acordo com as regras de oscilação (*wooble*), U na primeira posição do anticódon pode parear com A ou G na terceira posição do códon; então, existem dois códons que podem parear com esse anticódon:

Lembrete: Os códons são antiparalelos e complementares aos anticódons.

Dica: Consulte as regras de oscilação (*wooble*) no Quadro 15.2.

a. Anticódon: 5′-UUU-3′
Códon: 3′-AAA-5′
Códon: 3′-GAA-5′

Listando esses códons de maneira convencional com a extremidade 5′ à direita, temos:

Códon: 5′-AAA-3′

Códon: 5′-AAG-3′

De acordo com a Figura 15.8 , ambos códons especificam o aminoácido lisina (Lys). Lembre-se de que a oscilação (*wobble*) na terceira posição possilibita que mais de um códon especifique o mesmo aminoácido; então, qualquer oscilação que exista produzirá o mesmo aminoácido que os pareamentos de base padrão produziriam, e não precisamos calcular a oscilação para responder a esta pergunta. As respostas para as partes *b*, *c* e *d* são:

b.	Anticódon:	5′-GAC-3′
	Anticódon:	3′-CAG-5′
	Códon:	5′-GUC-3′ codifica Val
c.	Anticódon:	5′-UUG-3′
	Anticódon:	3′-GUU-5′
	Códon:	5′-CAA-3′ codifica Gln
d.	Anticódon:	5′-CAG-3′

Anticódon:	3′-GUC-5′
Códon:	5′-CUG-3′ codifica Leu

Questões de compreensão

Seção 15.1

1. O que é a hipótese de um gene, uma enzima? Por que essa hipótese foi um avanço importante em nossa compreensão de genética?

Seção 15.2

2. Quais são os métodos diferentes que foram usados para ajudar a desvendar o código genético? O que cada método revelou e quais são as vantagens e desvantagens de cada um?
3. O que são tRNAs isoaceptores?
4. Qual é a importância do fato de que muitos códons sinônimos diferem apenas na posição do terceiro nucleotídio?
5. Defina os termos a seguir à medida que eles se aplicam ao código genético:
   a. Quadro de leitura.
   b. Código sobreposto.
   c. Código não sobreposto.
   d. Códon de iniciação.
   e. Códon de terminação.
   f. Códon senso.
   g. Códon sem sentido.
   h. Código universal.
   i. Códons não universais.
6. Como é ajustado o quadro de leitura de um conjunto de sequência de nucleotídios?

Seção 15.3

7. Como os tRNAs são ligados a seus aminoácidos correspondentes?
8. Qual é o papel dos fatores de iniciação na síntese da proteína?
9. Quais são as diferenças do processo de iniciação nas células bacterianas e eucarióticas?
10. Descreva os fatores de alongamento usados na tradução bacteriana e explique o papel desempenhado por cada fator na tradução.
11. Quais eventos levam à terminação da tradução?
12. Compare e diferencie o processo de síntese proteica nas células bacterianas e eucarióticas, indicando as semelhanças e diferenças no processo de tradução nesses dois tipos de células.

Seção 15.4

13. Como as células procarióticas superam o problema de um ribossomo paralisado em um mRNA que não tem códon de terminação? Como as células eucarióticas solucionam esse mesmo problema?
14. Quais são os tipos de modificações pós-tradução das proteínas?
15. Explique como alguns antibióticos funcionam ao afetar o processo de síntese de proteínas.

Questões e problemas aplicados

Seção 15.1

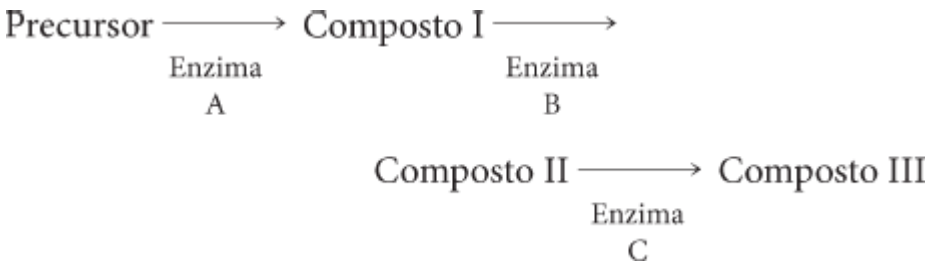
*16. Sydney Brenner isolou mutantes da *Salmonella typhimurium* que estavam envolvidos na biossíntese de triptofano e não cresceriam no meio mínimo. Quando esses mutantes foram testados no meio mínimo no qual um dos quatro compostos (fosfato de indol glicerol, indol, ácido antranílico e triptofano) foram adicionados, foram obtidas as respostas de crescimento apresentadas no quadro a seguir.

Meio	Ácido	Fosfato de indol

Mutante	mínimo	antranílico	glicerol	Indol	Triptofano
trp-1	−	−	−	−	+
trp-2	−	−	+	+	+
trp-3	−	−	−	+	+
trp-4	−	−	+	+	+
trp-6	−	−	−	−	+
trp-7	−	−	−	−	+
trp-8	−	+	+	+	+
trp-9	−	−	−	−	+
trp-10	−	−	−	−	+
trp-11	−	−	−	−	+

Apresente a ordem do fosfato de indol glicerol, indol, ácido antranílico e triptofano em uma via bioquímica que leve à síntese de triptofano. Indique qual etapa na via é afetada por cada uma das mutações.

17. Os compostos I, II e III estão na seguinte via bioquímica:



A mutação *a* inativa a enzima A, a mutação *b* inativa a enzima B e a mutação *c* inativa a enzima C. Os mutantes, cada um com um desses defeitos, foram testados em meio mínimo ao qual foi adicionado o composto I, II ou III. Preencha os resultados esperados desses testes ao colocar um sinal de mais (+) para crescimento ou um sinal de menos (−) para não crescimento no quadro abaixo.

	Meio mínimo ao qual foi adicionado		
Cepa com mutação	Composto I	Composto II	Composto III
a			
b			
c			

Seção 15.2

18. Uma geneticista fez o experimento descrito na **Figura 15.6**, mas desta vez ela combinou o nucleotídio guanina (em vez de uracila) com o polinucleotídio fosforilase. A proteína marcada com radioatividade aparecerá em qual tubo?
19. Para o experimento descrito na **Figura 15.6**, Nirenberg e Matthaei poderiam ter substituído a RNA polimerase em vez da polinucleotídio fosforilase sem modificar o experimento? Por que sim ou por que não?

*20. Considere que o número de diferentes tipos de base no RNA é quatro. Qual seria o tamanho mínimo do códon (número de nucleotídios) necessário para especificar todos os aminoácidos se o número de diferentes tipos de aminoácidos nas proteínas fosse: (a) 2, (b) 8, (c) 17, (d) 45, (e) 75?

*21. Quantos códons seriam possíveis em um código de trincas se fossem usadas apenas três bases (A, C e U)?

*22. Considerando o código genético apresentado na **Figura 15.8**, indique os aminoácidos especificados pelas seguintes sequências de mRNA bacterianas.
   a. 5′-AUGUUUAAAUUUAAAUUUUGA-3′.
   b. 5′-AGGGAAAUCAGAUGUAUAUAUAUAUAUGA-3′.
   c. 5′-UUUGGAUUGAGUGAAACGAUGGAUGAAAGAUUUCUCGCUUGA-3′.
   d. 5′-GUACUAAGGAGGUUGUAUGGGUUAGGGGCAUCAUUUUGA-3′.

23. Uma fita não molde no DNA bacteriano tem a sequência a seguir. Qual sequência de aminoácido será codificada por esta

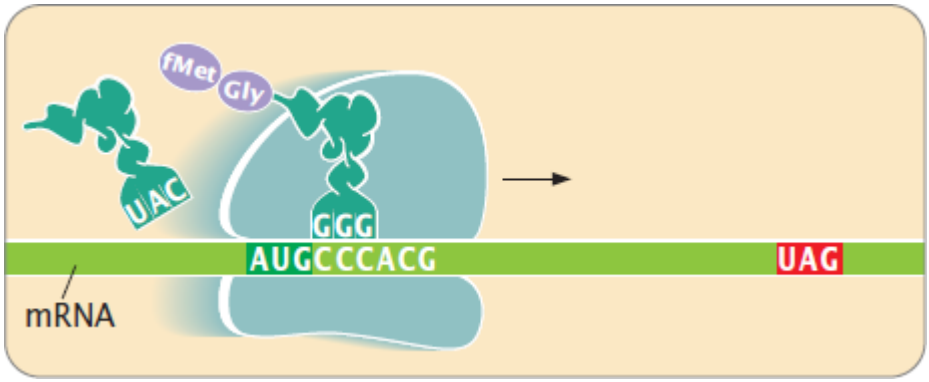
sequência?

5′-ATGATACTAAGGCCC-3′

24. A sequência de aminoácido a seguir é encontrada em um tripeptídio: Met-Trp-His. Indique todas as possíveis sequências de nucleotídios no mRNA, na fita molde de DNA e na fita não molde de DNA que podem codificar este tripeptídio.
25. Quantas sequências de mRNA diferentes podem codificar uma cadeia polipeptídica com a sequência de aminoácidos Met-Leu-Arg? (Certifique-se de incluir o códon de parada.)
*26. Vários tRNAs têm os seguintes anticódons. Considere as regras de *wobble* listadas no **Quadro 15.2** e forneça todos os possíveis códons com os quais cada tRNA pode parear.
    a. 5′-GGC-3′.
    b. 5′-AAG-3′.
    c. 5′-IAA-3′.
    d. 5′-UGG-3′.
    e. 5′-CAG-3′.
27. Um pesquisador criou copolímeros aleatórios de três nucleotídios ao misturar a polinucleotídio fosforilase com os nucleotídios guanina e adenina em uma razão de cinco nucleotídios guanina para um nucleotídio adenina. Forneça os diferentes copolímeros produzidos e suas proporções teóricas.
28. Considere que o nucleotídio na extremidade 5′ do primeiro anticódon do tRNA (o tRNA à esquerda) na Figura 15.9 foi mudado de G para U. Indique todos os códons com os quais o novo anticódon mutante pode parear.
29. Qual das mudanças de aminoácido a seguir poderia surgir a partir de uma mutação que mudou uma única base? Para cada mudança que possa surgir a partir da alteração de uma única base, determine qual posição do códon (primeiro, segundo ou terceiro nucleotídio) no mRNA deve ser alterada para que a mudança tenha resultado.
    a. Leu → Gln.
    b. Phe → Ser.
    c. Phe → Ile.
    d. Pro → Ala.
    e. Asn → Lys.
    f. Ile → Asn.

Seção 15.3

*30. Organize os componentes de tradução a seguir na ordem aproximada em que eles apareceriam ou seriam usados na síntese de proteína procariótica:
complexo de iniciação 70S
complexo de iniciação 30S
fator de liberação 1
fator de alongamento G
fator de iniciação 3
fator de alongamento Tu
fMet-tRNAf^{Met}
31. Examine a **Figura 15.12** de um tRNA. O que você acha que seria o efeito potencial de uma mutação na parte do gene do tRNA que codifica: (a) a haste aceptora; (b) o anticódon; (c) um dos nucleotídios corados de vermelho?
32. O diagrama a seguir ilustra uma etapa no processo de tradução. Esboce o diagrama e identifique os elementos:



a. As extremidades 5′ e 3′ do mRNA.
b. Os sítios A, P e E.
c. Códon de iniciação.
d. Códon de parada.
e. As extremidades amino e carboxila da cadeia de polipeptídios recém-sintetizada.
f. Localização aproximada da próxima ligação peptídica que será formada.
g. Local no ribossomo onde o fator de liberação 1 vai se ligar.

*33. Consulte o diagrama no Problema 32 para responder às perguntas a seguir.
a. Qual será o anticódon do próximo tRNA adicionado ao sítio A do ribossomo?
b. Qual será o próximo aminoácido adicionado à cadeia polipeptídica crescente?

*34. Um mRNA sintético adicionado a um sistema sintetizador de proteína fora da célula produz um peptídio com a seguinte sequência de aminoácidos: Met-Pro-Ile-Ser-Ala. Qual seria o efeito na tradução se os componentes a seguir fossem omitidos do sistema sintetizador de proteína fora da célula? Qual tipo de proteína seria produzido? Explique sua justificativa.
a. Fator de iniciação 3.
b. Fator de iniciação 2.
c. Fator de alongamento Tu.
d. Fator de alongamento G.
e. Fatores de liberação RF-1, RF-2 e RF-3.
f. ATP.
g. GTP.

35. Para cada uma das sequências a seguir, marque com um X o espaço correspondente ao processo *mais imediatamente* afetado ao eliminar a sequência. Escolha apenas um processo para cada sequência (marque um X por sequência).

	Processo mais imediatamente afetado após a deleção			
Sequência eliminada	**Replicação**	**Transcrição**	**Processamento de RNA**	**Tradução**
a. Sítio *ori*	_____	_____	_____	_____
b. Sítio consenso de *splicing* 3′	_____	_____	_____	_____
c. Cauda poli(A)	_____	_____	_____	_____
d. Terminador	_____	_____	_____	_____
e. Códon de iniciação	_____	_____	_____	_____
f. Consenso −10	_____	_____	_____	_____
g. Shine-Dalgarno	_____	_____	_____	_____

ANÁLISE DE DADOS

36. MicroRNAs são pequenas moléculas de RNA que se ligam à extremidade 3′ do mRNA e suprimem a tradução (ver *Capítulo 14*). Ainda está sendo investigado como os miRNAs suprimem a tradução. Alguns mRNAs eucarióticos têm sítios internos de ligação ao ribossomo *downstream* do cap 5′, onde os ribossomos normalmente se ligam. Em uma investigação, miRNAs não suprimiram a tradução dos ribossomos que se prendem aos sítios internos de ligação ao ribossomo (R. S. Pillai *et al*. 2005. *Science* 309:1573-1576). O que essa descoberta sugere sobre como os miRNAs suprimem a tradução?

37. Indique a sequência de aminoácidos da proteína codificada pelo mRNA na **Figura 15.19**.

Seção 15.4

*38. As mutações que introduzem códons de parada provocam várias doenças genéticas. Por exemplo, de 2 a 5% das pessoas com fibrose cística têm uma mutação que provoca um códon de parada prematuro no gene que codifica o regulador de condutância transmembrana na fibrose cística (CFTR). Esse códon de parada prematura produz uma forma truncada de CFTR não funcional e resulta nos sintomas da fibrose cística. Uma forma possível de tratar as pessoas com doenças genéticas causadas por estes tipos de mutações é enganar o ribossomo que faz a leitura através o códon de parada, inserindo um aminoácido em seu lugar. Embora a proteína produzida possa ter um aminoácido alterado, é mais provável que ela seja ao menos parcialmente funcional do que a proteína truncada produzida quando o ribossomo para no códon de parada. De fato, os geneticistas estão fazendo estudos clínicos em pessoas com fibrose cística usando um fármaco chamado PTC124, que interfere na capacidade do ribossomo em ler corretamente os códons de parada (C. Ainsworth. 2005. *Nature* 438:726-728). Com base no que sabemos sobre o mecanismo do decaimento do mRNA mediado por mutação sem sentido (NMD), você esperaria que o NMD seja um problema com este tipo de tratamento? Por que sim ou por que não?

Questões desafiadoras

Seção 15.2

39. A redundância do código genético significa que alguns aminoácidos podem ser especificados por mais de um códon. Por exemplo, o aminoácido leucina é codificado por seis códons diferentes. Dentro de um genoma, não há códons sinônimos na mesma quantidade; alguns aparecem com maior frequência que outros, e os códons preferidos são diferentes entre as espécies. Por exemplo, em uma espécie, o códon UUA pode ser usado mais frequentemente para codificar leucina, enquanto, em outras espécies, o códon CUU pode ser usado com maior frequência. Especule sobre um motivo para essa tendência no uso de códon e por que os códons preferidos não são os mesmos em todos os organismos.

Seção 15.3

40. Quais são as semelhanças entre os spliceossomos e os ribossomos? E as diferenças? Você pode sugerir alguns possíveis motivos para suas semelhanças?

*41. Foram feitos vários experimentos para obter informações sobre como o ribossomo eucariótico reconhece o códon de iniciação AUG. Em um experimento, o gene que codifica o tRNA iniciador da metionina ($tRNA_i^{Met}$) foi localizado e alterado. Os nucleotídios que especificam o anticódon $tRNA_i^{Met}$ foram alterados de modo que o anticódon no tRNA era 5′-CCA-3′ em vez de 5′-CAU-3′. Quando este gene mutante foi colocado em uma célula eucariótica, ocorreu a síntese proteica, mas as proteínas produzidas eram anormais. Algumas das proteínas produzidas tinham aminoácidos extras, e outras apresentavam alguns aminoácidos a menos que o normal.

a. O que esses resultados indicam sobre como o ribossomo reconhece o ponto de início da tradução nas células eucarióticas? Explique sua justificativa.
b. Se o mesmo experimento for feito em células bacterianas, que resultados você esperaria?
c. Explique por que algumas das proteínas tinham aminoácidos extras e outras tinham alguns aminoácidos a menos que o normal.

16 Controle da Expressão Gênica nas Bactérias

Óperons e a célula de ruído

Em 2011, geneticistas de todo o mundo celebraram o aniversário de 50 anos do óperon. Óperon é um grupo de genes que compartilham o mesmo promotor e são transcritos como uma unidade, produzindo um único mRNA que codifica várias proteínas. Tipicamente essas proteínas interagem de alguma forma; por exemplo, o óperon *trp* na bactéria *Escherichia coli* codifica componentes das três enzimas que interagem para sintetizar o aminoácido triptofano. Os óperons controlam a expressão de genes como esses ao regular sua transcrição. Os genes nas bactérias são organizados em óperons, mas eles são muito menos comuns nos eucariotos.

O óperon foi descoberto em uma excelente pesquisa de François Jacob e Jacques Monod, que trabalharam em extremidades opostas do sótão do Instituto Pasteur, em Paris. Jacob estava estudando o bacteriófago lambda, um vírus que infecta a *E. coli*; Monod estava analisando as propriedades da *β*-galactosidase, uma enzima que a *E. coli* usa para metabolizar o açúcar lactose. Em uma noite do verão de 1958, Jacob teve uma inspiração – ele vislumbrou uma conexão entre a pesquisa em andamento em cada extremidade do sótão. Jacob identificou que os genes que induziam a reprodução do fago eram controlados da mesma maneira que os genes que controlam a produção da *β*-galactosidase na *E. coli.* Essa descoberta levou a uma importante colaboração entre Jacob e Monod; juntos, eles acabaram descobrindo a estrutura e a função do óperon *lac* e, em 1965, receberam o Prêmio Nobel em Fisiologia ou Medicina junto com seu colaborador Andre Lwoff.

Após o achado de Jacob e Monod do óperon *lac* foram descobertos outros óperon nas bactérias, e uma grande parcela dos estudos se voltou para o mecanismo da função do óperon. Apesar da abrangente pesquisa sobre *como* os óperons trabalham, muito pouco se sabia sobre o *porquê* da existência dos óperons: por que os procariotos têm óperons enquanto a maioria dos eucariotos não tem? Por que alguns genes são incluídos nos óperons e outros não? Essas perguntas fascinaram Oleg Igoshin, da Universidade de Rice, e Christian Ray, da Universidade de Texas MD Anderson Cancer Center. Igoshin e Ray são biólogos computacionais, um novo ramo de cientistas que usam matemática complexa para estudar problemas fundamentais de biologia. Eles adotaram a uma abordagem que parecia ser improvável para a pergunta sobre a existência dos óperons. Em vez de usar as bactérias em crescimento, induzir mutações e examinar o DNA, eles desenvolveram vários modelos matemáticos de redes de genes que poderiam ser trabalhados no computador. Com esses modelos, eles observaram como os genes funcionavam quando agrupados em óperon e quando regulados separadamente.

Igoshin e Ray sabiam que ocorriam flutuações aleatórias naturais nos níveis de transcrição e tradução. Por causa dessas flutuações – ruídos no sistema –, as quantidades de diferentes proteínas podem variar muito; a quantidade de proteína produzida pode ser superior ou inferior ao que seria ideal para o crescimento e a sobrevivência da célula. Igoshin e Ray desenvolveram uma hipótese de que coordenar a transcrição de vários genes por meio de uma estrutura de óperon poderia reduzir o ruído no sistema e possibilitaria um controle mais acurado sobre a expressão gênica.

Para testar sua hipótese, eles executaram modelos computacionais para seis diferentes tipos de interações dos produtos dos genes potencialmente encontrados nos óperons. Seus modelos mostraram que, para alguns tipos de interações de proteína, agrupar os genes em um óperon reduzia o ruído bioquímico. Para outros tipos de interações de proteínas, agrupar os genes em óperons de fato aumentou o ruído. Assim, a estrutura do óperon tem o potencial de aumentar ou reduzir o ruído, dependendo de quais genes são agrupados.

Igoshin e Ray examinaram os genes realmente encontrados em óperons na *E. coli* e descobriram que os óperons com genes cujas interações reduziam o ruído eram mais comuns do que o esperado aleatoriamente. Por outro lado, aqueles cujas interações gênicas aumentaram o ruído eram menos comuns do que o esperado. Eles concluíram que os óperons evoluíram como uma maneira para a célula encaixar a transcrição dos genes para reduzir o ruído bioquímico, fazendo com que a célula ajuste mais sutilmente as proporções relativas de proteínas codificadas pelo óperon. Os cientistas especularam que os óperons seriam menos comuns nos eucariotos porque o volume celular maior reduz o efeito das flutuações aleatórias e, talvez, porque os eucariotos

tenham outros mecanismos (como mudanças na estrutura da cromatina) que encaixam a transcrição dos genes.

Este é o primeiro de dois capítulos sobre **regulação gênica**, os mecanismos e os sistemas que controlam a expressão dos genes. Neste capítulo, vamos analisar os sistemas de regulação gênica nas bactérias. Começamos analisando a necessidade da regulação gênica, os níveis em que a expressão gênica é controlada e a diferença entre genes e elementos reguladores. Então, examinaremos a estrutura e a função dos óperons e analisaremos a regulação gênica em alguns exemplos específicos de óperons, incluindo o óperon *lac* estudado por Jacob e Monod. Por fim, discutiremos vários tipos de regulação gênica bacteriana facilitados pelas moléculas de RNA. O Capítulo 17 descreve os mecanismos de regulação gênica nos genomas eucarióticos.

16.1 A regulação da expressão gênica é crítica para todos os organismos

Um tema importante da genética molecular é o dogma central, que afirma que a informação genética flui do DNA para o RNA e para as proteínas (ver Figura 10.16). Embora o dogma central tenha fornecido uma base molecular para a conexão entre o genótipo e o fenótipo, ele não conseguiu responder a uma pergunta crítica: como é regulado o fluxo de informação ao longo da via molecular?

Considere a *E. coli*, uma bactéria encontrada no intestino grosso. Os hábitos alimentares determinam os nutrientes disponíveis para esta bactéria: ela não pode procurar alimentos quando os nutrientes são escassos nem se deslocar quando o ambiente é desfavorável. A *E. coli* compensa sua incapacidade de alterar o ambiente externo ao ser internamente flexível. Por exemplo, se houver glicose no ambiente, ela usa esse açúcar para gerar ATP; na ausência de glicose, ela usa lactose, arabinose, maltose, xilose ou qualquer um dos vários açúcares. Quando há aminoácidos, a *E. coli* os usa para sintetizar proteínas; na ausência de um aminoácido específico, a bactéria produz as enzimas necessárias para sintetizá-los. Assim, a *E. coli* responde às mudanças ambientais ao alterar rapidamente sua bioquímica. Essa flexibilidade bioquímica, entretanto, tem um preço alto. Produzir todas as enzimas necessárias para cada quadro ambiental consumiria muita energia. Então, como a *E. coli* mantém a flexibilidade bioquímica enquanto otimiza a eficiência energética?

A resposta é a regulação gênica. As bactérias carreiam informações genéticas para sintetizar muitas proteínas, mas apenas um subconjunto dessa informação genética é expresso. Quando o ambiente muda, novos genes são expressos e as proteínas adequadas para o novo ambiente são sintetizadas. Por exemplo, se uma fonte de carbono surge no ambiente, genes codificadores de enzimas que captam e metabolizam essa fonte de carbono são rapidamente transcritos e traduzidos. Quando essa fonte de carbono desaparece, esses genes são “desligados”.

Os organismos eucarióticos multicelulares encaram um dilema diferente. Células individuais em um organismo multicelular são especializadas em tarefas específicas. As proteínas produzidas por um neurônio, por exemplo, são bem diferentes das produzidas por um leucócito. Embora tenham formato e função diferentes, um neurônio e um leucócito ainda têm as mesmas instruções genéticas.

O desafio do organismo multicelular é obter a especialização das células que têm um mesmo conjunto de instruções genéticas (o processo de desenvolvimento). Esse desafio é solucionado pela regulação gênica: todas as células de um organismo carreiam a mesma informação genética, mas apenas um subconjunto de genes é expresso em cada tipo celular. Os genes necessários para os outros tipos de células não são expressos. A regulação gênica é, portanto, crucial para a flexibilidade unicelular e a especialização multicelular, e é crítica para o sucesso de todos os organismos vivos.

Conceitos

Nas bactérias, a regulação gênica mantém a flexibilidade interna, ligando e desligando os genes em resposta às mudanças ambientais. Nos organismos multicelulares, a regulação gênica produz a diferenciação celular.

Os mecanismos de regulação gênica foram investigados pela primeira vez nas células bacterianas, nas quais a disponibilidade de mutantes e a facilidade de manipulação laboratorial possibilitaram desvendar os mecanismos. Quando começou o estudo desses mecanismos nas células eucarióticas, a regulação gênica bacteriana parecia ser muito diferente da regulação gênica eucariótica. Entretanto, à medida que mais informação foi acumulada sobre a regulação gênica, surgiram vários itens em comum. Atualmente, constatou-se que muitos aspectos da regulação gênica nas células bacterianas e eucarióticas são semelhantes. Antes de examinar elementos específicos da regulação gênica bacteriana (tema deste capítulo) e a regulação gênica eucariótica (no *Capítulo 17*), analisaremos rapidamente alguns itens da regulação gênica comuns a todos os organismos.

Genes e elementos regulatórios

Ao analisar a regulação gênica nas bactérias e nos eucariotos, é crucial diferenciar as sequências de DNA que são transcritas daquelas que regulam a expressão de outras sequências. Os **genes estruturais** codificam as proteínas que atuam no metabolismo ou na biossíntese ou têm uma função estrutural na célula. Os **genes regulatórios** são aqueles cujos produtos (RNA ou proteínas) interagem com outras sequências de DNA e afetam a transcrição ou a tradução dessas sequências. Em muitos casos, os produtos dos genes regulatórios são proteínas ligadoras de DNA (embora as moléculas de RNA também afetem a expressão gênica). As bactérias e os eucariotos usam genes regulatórios para controlar a expressão de muitos dos seus genes estruturais. Entretanto, alguns genes estruturais, sobretudo os que codificam funções celulares essenciais, são expressos de forma contínua e são chamados de **genes constitutivos**. Portanto, os genes constitutivos não são regulados.

Também encontraremos sequências de DNA que não são transcritas de forma alguma, mas ainda participam na regulação dos genes e de outras sequências de DNA. Esses **elementos reguladores** afetam a expressão de sequências às quais estão fisicamente ligados. Os elementos reguladores são comuns nas células bacterianas e eucarióticas, e boa parte da regulação gênica em ambos os tipos de organismos se deve à ação de proteínas produzidas por genes regulatórios que reconhecem e se ligam a elementos reguladores.

A regulação da expressão gênica pode ocorrer por processos que estimulam a expressão gênica, chamados de *controle positivo*, ou por meio de processos que inibem a expressão gênica, chamados de *controle negativo*. As bactérias e os eucariotos usam os mecanismos de controle positivo e negativo para regular seus genes. Entretanto, o controle negativo é mais importante nas bactérias, enquanto é mais provável que os eucariotos usem os mecanismos de controle positivo.

Conceitos

Os elementos reguladores são sequências de DNA que não são transcritas, mas afetam a expressão dos genes. Os mecanismos de controle positivo estimulam a expressão gênica, enquanto o controle negativo inibe a expressão gênica.

Checagem dos conceitos 1

O que é um gene constitutivo?

Níveis de regulação gênica

Nas bactérias e nos eucariotos, os genes podem ser regulados em vários níveis ao longo da via de fluxo de informação do genótipo para o fenótipo (**Figura 16.1**). Primeiro, a regulação pode ser feita por meio da alteração do DNA ou da estrutura da cromatina; esse tipo de regulação gênica ocorre principalmente nos eucariotos. As modificações no DNA ou no seu acondicionamento podem ajudar a determinar quais sequências estão disponíveis para transcrição ou a taxa de transcrição das sequências. A metilação do DNA e as alterações na cromatina são dois processos cruciais na regulação gênica.

Um segundo ponto no qual um gene pode ser regulado é no seu nível de transcrição. Para fins de economia da célula, limitar a produção de uma proteína no início do processo faz sentido, e a transcrição é um ponto importante da regulação gênica nas células bacterianas e eucarióticas. Um terceiro ponto potencial da regulação gênica é o processamento do mRNA. O mRNA eucariótico é muito modificado antes de ser traduzido: é adicionado um cap 5′, a extremidade 3′ é clivada, e poliadenilada e íntrons são removidos (ver *Capítulo 14*). Essas modificações determinam a estabilidade do mRNA, o movimento do mRNA no citoplasma, se o mRNA pode ser traduzido, a taxa de tradução e a sequência de aminoácidos da proteína produzida. Há evidências crescentes de que vários mecanismos regulatórios nas células eucarióticas operem no nível do processamento do mRNA.

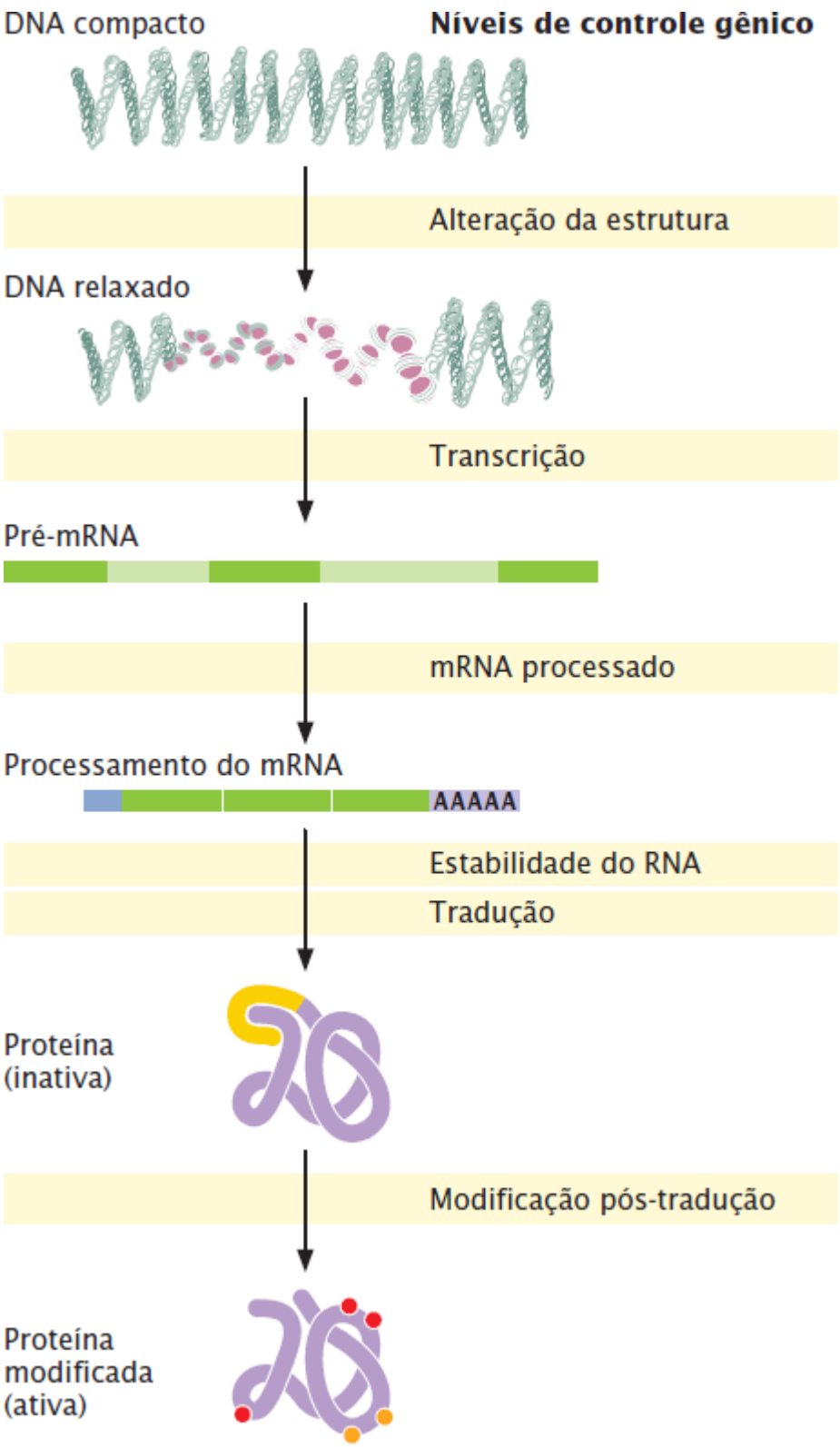


Figura 16.1 **A expressão gênica pode ser controlada em múltiplos níveis.**

Um quarto ponto para o controle da expressão gênica é a regulação da estabilidade do RNA. A quantidade de proteína produzida depende não apenas da quantidade de mRNA sintetizado, mas também da taxa de degradação do mRNA. Um quinto ponto da regulação gênica é no nível de tradução, um processo complexo que requer muitas enzimas, fatores de proteínas e moléculas de RNA (ver *Capítulo 15*). Todos esses fatores, assim como a disponibilidade dos aminoácidos, influenciam a taxa de produção das proteínas e, portanto, fornecem pontos nos quais a expressão gênica pode ser controlada. A tradução também pode ser afetada por sequências no mRNA.

Finalmente, muitas proteínas são modificadas após a tradução (ver *Capítulo 15*), e essas modificações afetam se as proteínas se tornam ativas; os genes podem ser regulados por meio de processos que afetam a modificação pós-tradução. A expressão gênica pode ser afetada por atividades regulatórias em qualquer um ou em todos esses pontos.

Conceitos

A expressão gênica pode ser controlada em qualquer um de vários níveis ao longo da via molecular do DNA para a proteína, incluindo a estrutura do DNA ou da cromatina, transcrição, processamento do mRNA, estabilidade do RNA, tradução e modificação pós-tradução.

Checagem dos conceitos 2

Por que a transcrição é um nível particularmente importante da regulação gênica nas bactérias e nos eucariotos?

Proteínas ligadoras do DNA

A maior parte da regulação gênica nas bactérias e nos eucariotos é feita por proteínas que se ligam a sequências de DNA e afetam sua expressão. Essas proteínas reguladoras têm partes funcionais bem definidas, chamadas **domínios**, compostas tipicamente por 60 a 90 aminoácidos, responsáveis pela ligação ao DNA. Dentro de um domínio, apenas alguns aminoácidos de fato entram em contato com o DNA. Esses aminoácidos (mais comumente asparagina, glutamina, glicina, lisina e arginina) formam pontes de hidrogênio com as bases ou interagem com o esqueleto de açúcar-fosfato do DNA. Muitas proteínas reguladoras têm domínios adicionais que se ligam a outras moléculas como outras proteínas reguladoras. Ao se ligar fisicamente ao DNA, essas proteínas podem afetar a expressão de um gene. A maioria das proteínas ligadoras de DNA se liga de maneira dinâmica, o que significa que elas se unem e se separam temporariamente do DNA e de outras proteínas reguladoras. Assim, embora muitas gastem a maior parte do seu tempo atadas ao DNA, essa ligação nunca é permanente. Essa natureza dinâmica significa que outras moléculas podem competir com essas proteínas pelos sítios regulatórios do DNA.

Essas proteínas podem ser agrupadas em muitos tipos distintos com base na estrutura característica, chamada de motivo, encontrada no domínio ligador. Os motivos são estruturas simples, como as alfa-hélices, que podem se ajustar no sulco principal do DNA. Por exemplo, o motivo hélice-giro-hélice (**Figura 16.2 A**), composto por duas alfa-hélices conectadas por um giro, é comum nas proteínas regulatórias bacterianas. O motivo dedo-de-zinco (**Figura 16.2 B**), comum na maioria das proteínas regulatórias eucarióticas, é composto por uma alça de aminoácidos contendo um íon zinco. O zíper de leucina (**Figura 16.2 C**) é outro motivo encontrado em várias proteínas reguladoras eucarióticas. O **Quadro 16.1** apresenta esses motivos comuns de ligação ao DNA e outros motivos.

Conceitos

As proteínas reguladoras que se ligam ao DNA têm motivos comuns que interagem com sequências no DNA.

Checagem dos conceitos 3

Como os aminoácidos nas proteínas ligadoras de DNA interagem com o DNA?

a. Ao formar ligações covalentes com bases de DNA.
b. Ao formar pontes de hidrogênio com bases de DNA.
c. Ao formar ligações covalentes com açúcares de DNA.

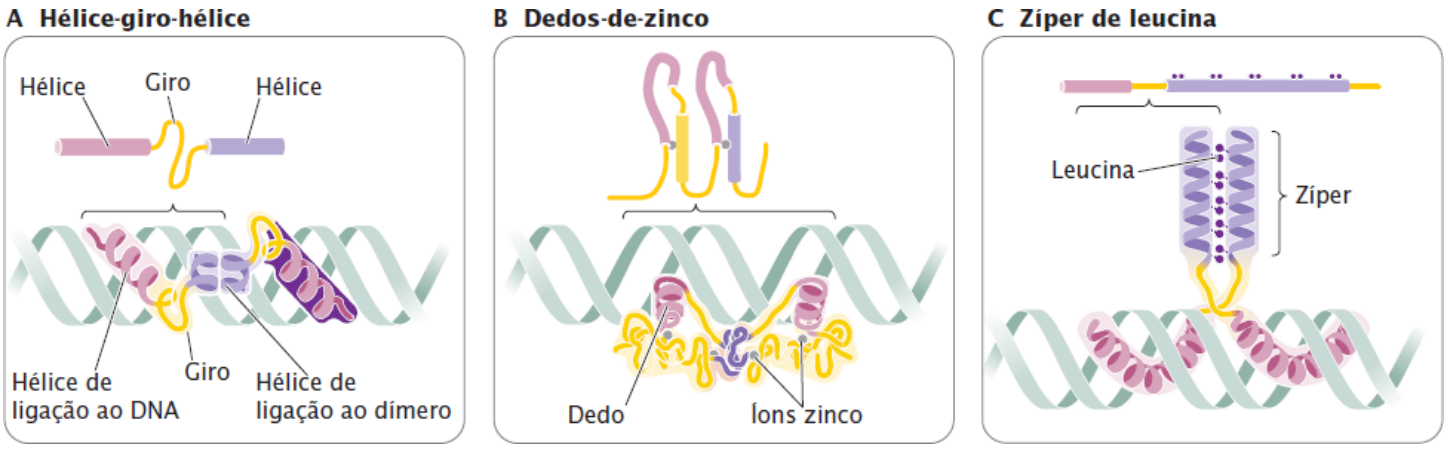


Figura 16.2 **As proteínas ligadoras de DNA podem ser agrupadas em vários tipos com base nas suas estruturas, ou motivos.**

Quadro 16.1 Motivos comuns de ligação ao DNA.

Motivo	Localização	Características	Sítio de ligação no DNA
Hélice-giro-hélice	Proteínas reguladoras bacterianas, motivos relacionados nas proteínas eucarióticas	Duas alfa-hélices	Sulco principal
Dedo-de-zinco	Proteínas reguladoras e outras proteínas	Alça de aminoácidos com o zinco na base	Sulco principal
Receptor de esteroide	Proteínas eucarióticas	Duas alfa-hélices perpendiculares com zinco circundadas por quatro resíduos de cistina	Sulco principal e esqueleto do DNA
Zíper de leucina	Fatores de transcrição eucarióticos	Hélice de resíduos de leucina e um braço básico; dois resíduos de leucina se interdigitam	Dois sulcos principais adjacentes
Hélice-giro-hélice	Proteínas eucarióticas	Duas alfa-hélices separadas por uma alça de aminoácidos	Sulco principal
Homeodomínio	Proteínas reguladoras eucarióticas	Três alfa-hélices	Sulco principal

16.2 Os óperons controlam a transcrição nas células bacterianas

Uma diferença significativa entre os genes bacteriano e eucariótico está na organização dos genes funcionais. Conforme discutido na introdução deste capítulo, muitos genes bacterianos com funções relacionadas estão agrupados e sob o controle de um único promotor. Esses genes são geralmente transcritos juntos em um único mRNA. Um grupo de genes estruturais bacterianos que são transcritos juntos (também com suas sequências promotoras e adicionais que controlam a transcrição) é chamado de **óperon**. O óperon regula a expressão dos genes estruturais ao controlar a transcrição, que nas bactérias é o nível mais importante da regulação gênica.

Estrutura do óperon

A **Figura 16.3** apresenta a organização de um óperon típico. Em uma extremidade do óperon existe um conjunto de genes estruturais, apresentados na Figura 16.3 como gene *a*, gene *b* e gene *c*. Esses genes estruturais são transcritos em um único mRNA, que é traduzido para produzir as enzimas A, B e C. Essas enzimas executam várias reações bioquímicas que convertem a molécula X precursora no produto Y. A transcrição dos genes estruturais *a*, *b* e *c* está sob controle de um promotor, que se encontra *upstream* do primeiro gene estrutural. A RNA polimerase se liga ao promotor e então se move *downstream*, transcrevendo os genes estruturais.

Um **gene regulador** ajuda a controlar a transcrição dos genes estruturais do óperon. Embora ele afete a função do óperon, esse gene não é considerado parte do óperon. O gene regulador tem seu próprio promotor e é transcrito em um mRNA curto, que é traduzido em uma pequena proteína. Essa **proteína reguladora** consegue se ligar a uma região do óperon chamada **operador**, e define se pode ocorrer a transcrição. Em geral, o operador se sobrepõe à extremidade 3′ do promotor e às vezes à extremidade 5′ do primeiro gene estrutural (ver Figura 16.3).

Conceitos

Os genes funcionais relacionados nas células bacterianas são agrupados como uma unidade de transcrição única chamada óperon. Um óperon típico inclui vários genes estruturais, um promotor para os genes

estruturais e um sítio operador ao qual o produto do gene regulador se liga.

Checagem dos conceitos 4

Qual é a diferença entre um gene estrutural e um gene regulador?

a. Os genes estruturais são transcritos em mRNA, mas os genes reguladores não.
b. Os genes estruturais têm estruturas complexas; os genes reguladores têm estruturas simples.
c. Os genes estruturais codificam proteínas que atuam na estrutura da célula; os genes reguladores são responsáveis pelas reações metabólicas.
d. Os genes estruturais codificam as proteínas; os genes reguladores controlam a transcrição dos genes estruturais.

Controles negativo e positivo | Óperons induzíveis e reprimíveis

Existem dois tipos de controle de transcrição: o **controle negativo**, no qual uma proteína reguladora é repressora, ligando-se ao DNA e inibindo a transcrição; e o **controle positivo**, no qual uma proteína reguladora é uma ativadora, estimulando a transcrição. Os óperons também podem ser induzíveis ou reprimíveis. Os **óperons induzíveis** são aqueles em que a transcrição está normalmente desligada (não ocorre); algo precisa acontecer para induzir a transcrição, ou ligá-la. Os **óperons reprimíveis** são aqueles em que a transcrição está normalmente ligada (ocorre); algo deve acontecer para reprimir a transcrição, ou desligá-la. Vamos analisar vários tipos de mecanismos básicos de controle nas seções a seguir.

Óperons induzíveis negativos. Em um óperon induzível negativo, o gene regulador codifica um *repressor* ativo que se liga com facilidade ao operador (**Figura 16.4 A**). Como o sítio do operador se sobrepõe ao sítio do promotor, a ligação dessa proteína ao operador bloqueia fisicamente a ligação da RNA polimerase ao promotor e impede a transcrição. Para que a transcrição ocorra, algo deve intervir para impedir a ligação do repressor no sítio do operador. Esse tipo de sistema é chamado de *induzível*, porque a transcrição está normalmente desligada (inibida) e tem de ser ligada (induzida).

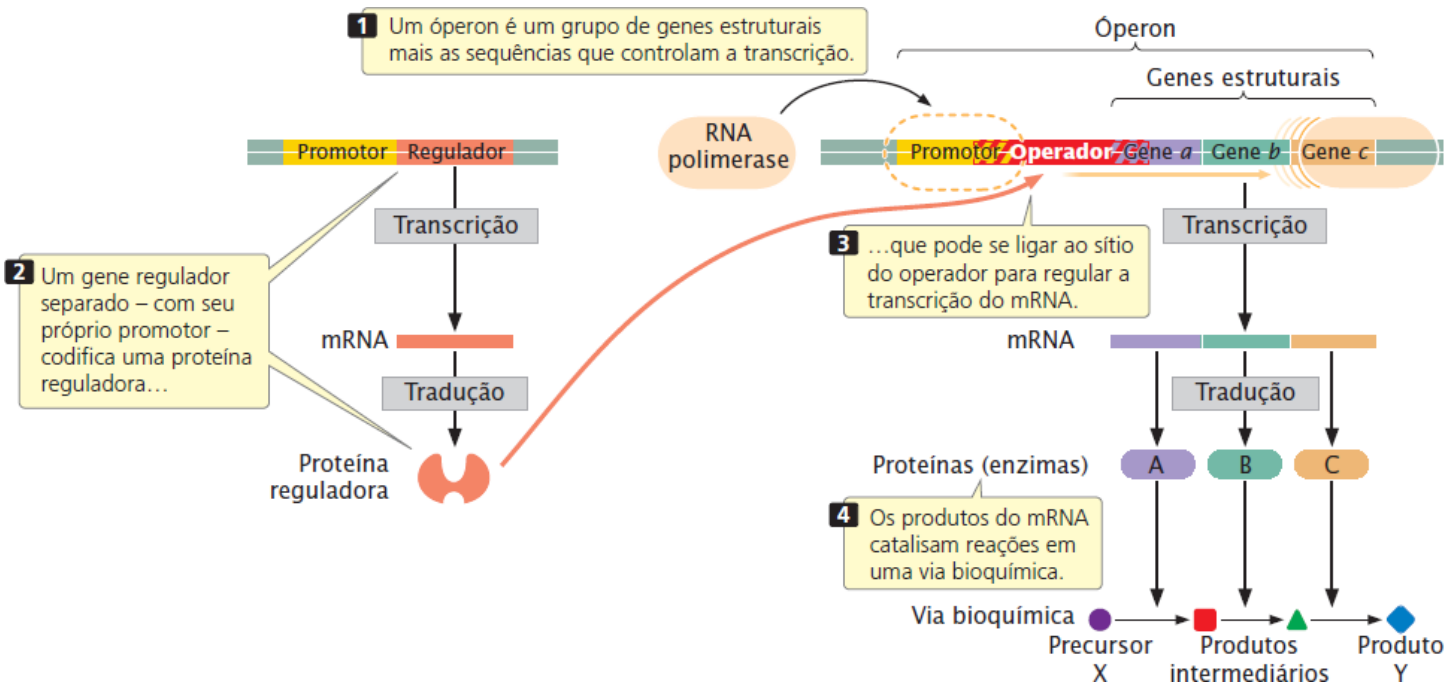


Figura 16.3 Um óperon é uma unidade de transcrição única que inclui uma série de genes estruturais, um promotor e um operador.

A transcrição é ligada quando uma pequena molécula, um **indutor**, liga-se ao repressor (**Figura 16.4 B**). As proteínas reguladoras têm frequentemente dois sítios de ligação: um que se liga ao DNA e outro que se liga a uma pequena molécula como um indutor. A ligação do indutor (precursor V na Figura 16.4 B) altera o formato do repressor, evitando que ele se ligue ao DNA. As proteínas desse tipo, que alteram o formato ao se ligar a outra molécula, são chamadas de **proteínas alostéricas**.

Quando o indutor está ausente, o repressor se liga ao operador, os genes estruturais não são transcritos e as enzimas D, E e F (que metabolizam o precursor V) não são sintetizadas (ver Figura 16.4 A). Esse mecanismo é adaptável: como o precursor V não está disponível, a síntese das enzimas seria um desperdício quando elas não têm substrato para metabolizar. À medida que o

precursor V se torna disponível, alguns se ligam ao repressor, deixando o repressor inativo e incapaz de se ligar ao sítio do operador. A RNA polimerase agora pode se ligar ao promotor e transcrever os genes estruturais. O mRNA resultante é então traduzido nas enzimas D, E e F, que convertem o substrato V no produto W (ver Figura 16.4 B). Então, um óperon com controle induzível negativo regula a síntese das enzimas de forma econômica: as enzimas são sintetizadas apenas quando seu substrato (V) está disponível.

Os óperons induzíveis controlam as proteínas que executam processos de degradação – as proteínas que degradam moléculas. Para esses tipos de proteínas, o controle induzível faz sentido porque as proteínas não são necessárias, a não ser que haja substrato (que é degradado pelas proteínas).

Óperons reprimíveis negativos. Alguns óperons com controle negativo são *reprimíveis*, o significa que a transcrição ocorre *normalmente* e tem de ser desligada, ou reprimida. A proteína reguladora nesse tipo de óperon também é uma repressora, mas é sintetizada em uma forma *inativa* que não pode se ligar por si só ao operador. Como nenhum repressor se liga ao operador, a RNA polimerase se liga com facilidade ao promotor e ocorre a transcrição dos genes estruturais (**Figura 16.5 A**).

Algo precisa acontecer para ativar o repressor e desligar a transcrição. Uma pequena molécula chamada de **correpressor** liga-se ao repressor e o torna capaz de se ligar ao operador. No exemplo ilustrado (ver Figura 16.5 A), o produto (U) da reação metabólica é o correpressor. Como o nível do produto U está elevado, ele está disponível para se ligar ao repressor e ativá-lo, evitando a transcrição (**Figura 16.5 B**). Com o óperon reprimido, as enzimas G, H e I não são sintetizadas e não é mais produzido U a partir de T. Entretanto, quando todo o produto U for usado, o repressor não é mais ativado pelo produto U e não pode se ligar ao operador. A inativação do repressor possibilita a transcrição dos genes estruturais e a síntese das enzimas G, H e I, resultando na conversão do precursor T no produto U. Como os óperons induzíveis, os óperons reprimíveis são econômicos: as enzimas são sintetizadas apenas se forem necessárias.

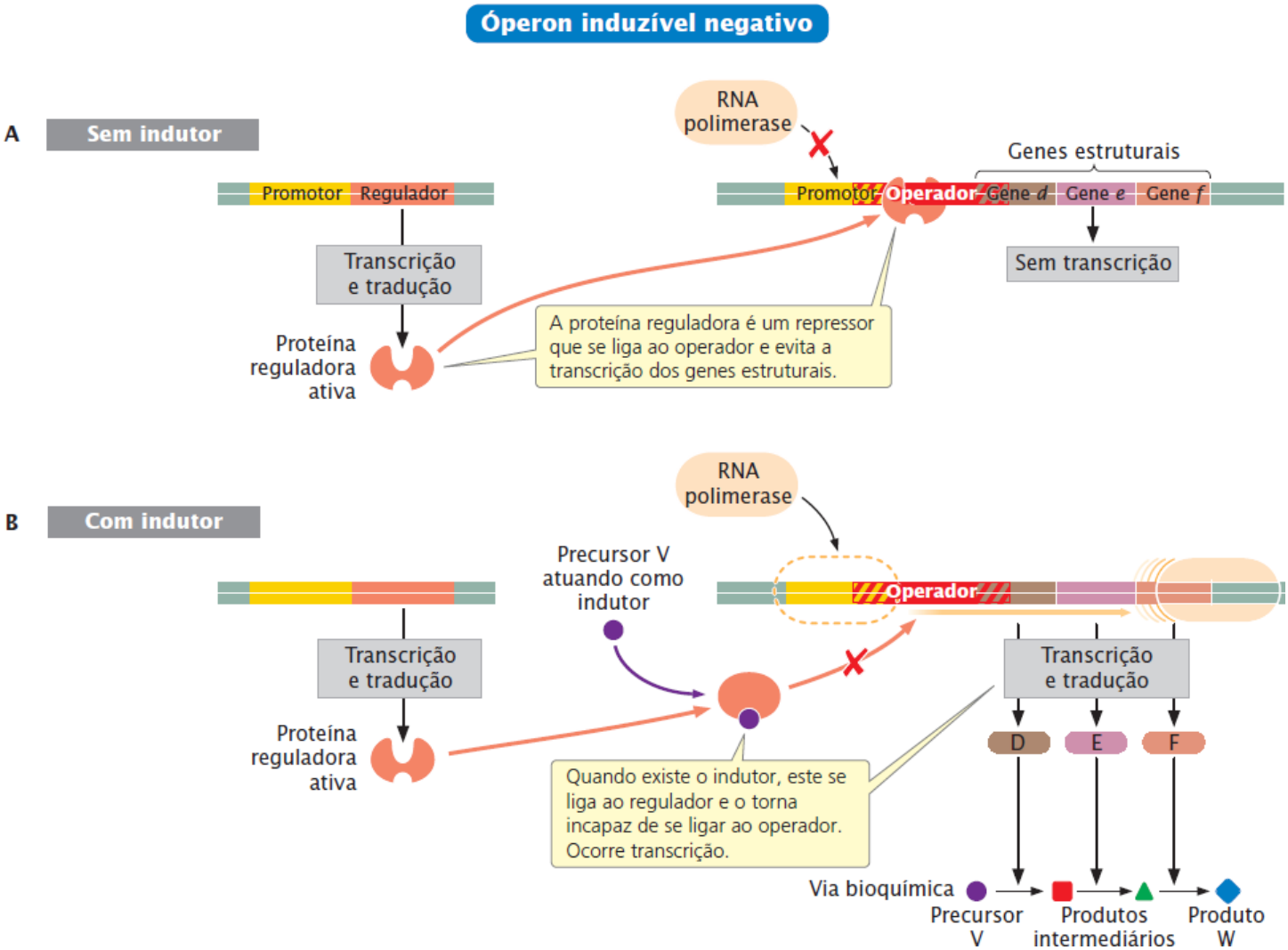


Figura 16.4 Alguns óperons são induzíveis.

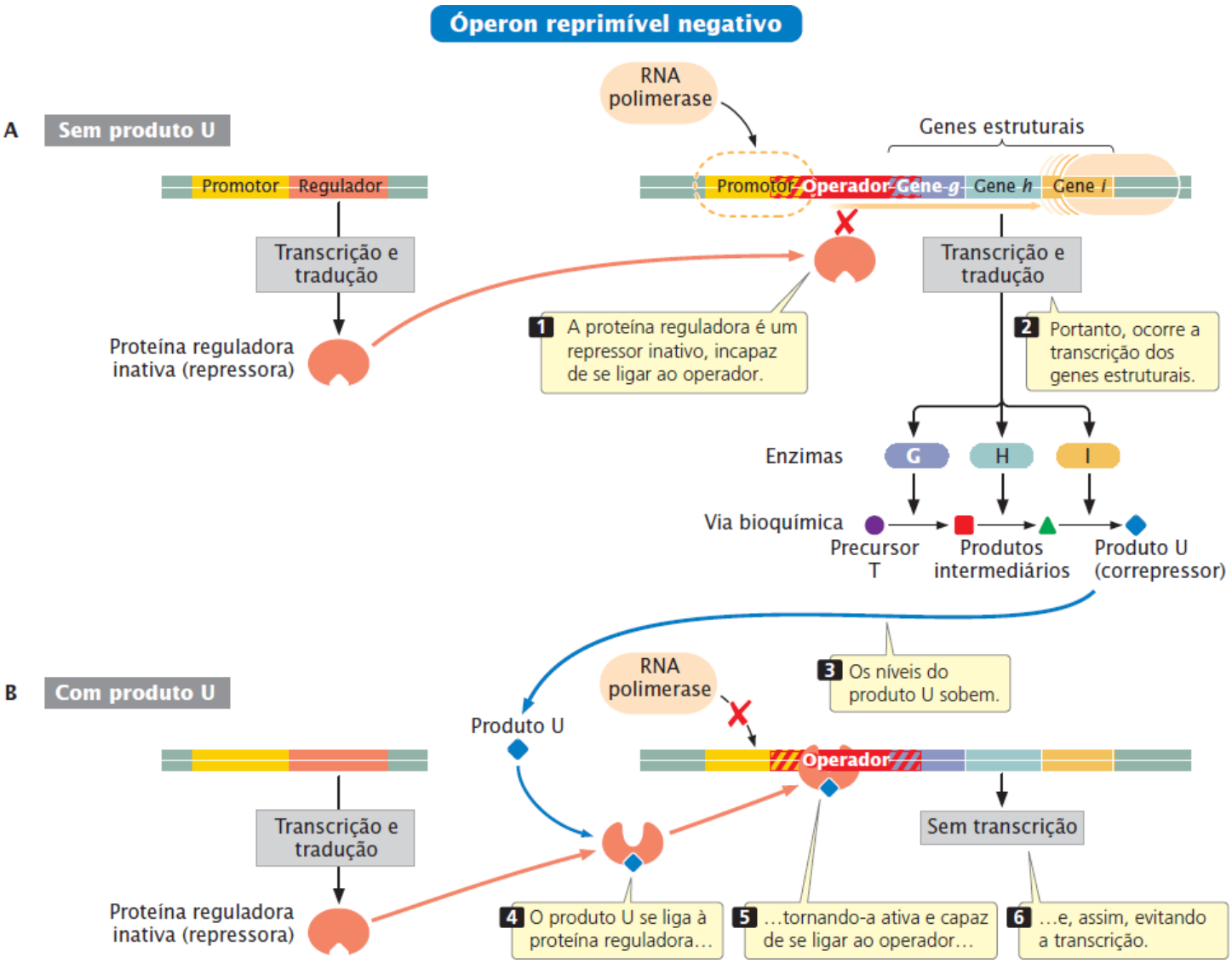


Figura 16.5 **Alguns óperons são reprimíveis.**

Os óperons reprimíveis controlam as proteínas responsáveis pela biossíntese das moléculas necessárias na célula, como os aminoácidos. Para esses tipos de óperons, o controle reprimível faz sentido porque a célula sempre precisa do produto produzido pela proteína. Assim, esses óperons estão normalmente ligados e são desligados quando há quantidades adequadas do produto já existente.

Observe que os sistemas induzíveis e reprimíveis que analisamos são formas de controle negativo, no qual a proteína reguladora é um repressor. Agora vamos analisar o controle positivo, no qual uma proteína reguladora estimula a transcrição.

Controle positivo. Com o controle positivo, uma proteína reguladora é um ativador: ela se liga ao DNA (em geral em um sítio diferente do operador) e estimula a transcrição. O controle positivo pode ser induzível ou reprimível.

Em um óperon induzível positivo, a transcrição está desligada porque a proteína reguladora (uma ativadora) é produzida na forma inativa. A transcrição ocorre quando um indutor se liga à proteína reguladora, liberando o regulador ativo. Do ponto de vista lógico, o indutor será o precursor da reação controlada pelo óperon de modo que as enzimas necessárias seriam sintetizadas apenas quando houvesse o substrato para sua reação.

Um óperon positivo também pode ser reprimível; a proteína reguladora é produzida em uma forma que se liga com facilidade ao DNA, o que significa que a transcrição ocorre normalmente e tem de ser reprimida. A transcrição é inibida quando um substrato se fixa ao ativador e o deixa incapaz de se ligar ao DNA de modo que a transcrição não é mais estimulada. Aqui, o produto (P) da reação controlada pelo óperon seria logicamente a substância reprimida, porque ela seria econômica para a célula para evitar a transcrição dos genes que possibilitam a síntese de P quando já existe produto P. As características do controle positivo e negativo nos óperons induzíveis e reprimíveis estão resumidas na **Figura 16.6**. Resolva o Problema 11

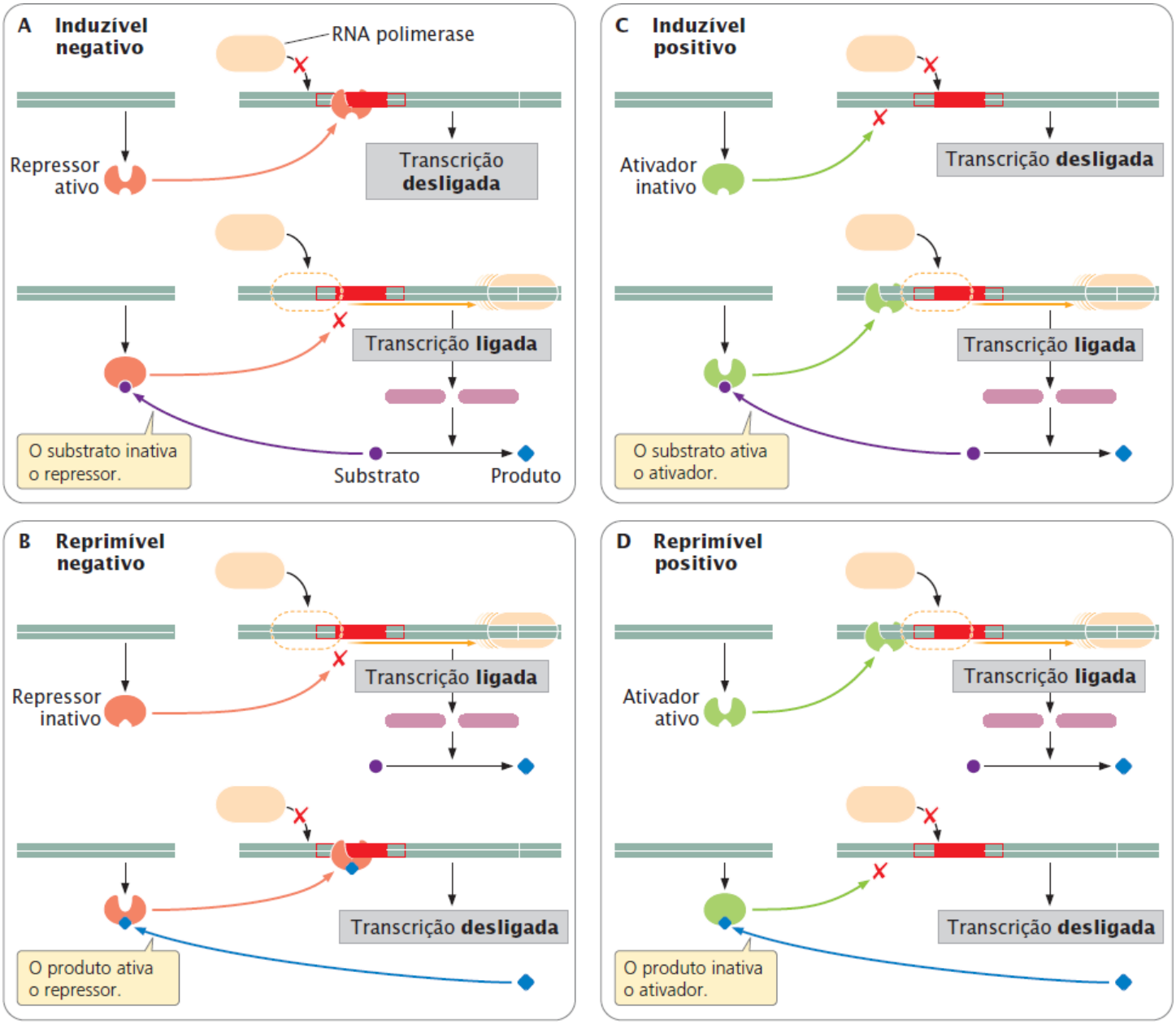


Figura 16.6 **Um resumo das características do controle positivo e negativo nos óperons induzíveis e reprimíveis.**

Conceitos

Existem dois tipos básicos de controle de transcrição: negativo e positivo. No controle negativo, quando uma proteína regulatória (repressora) se liga ao DNA, a transcrição é inibida. No controle positivo, quando uma proteína regulatória (ativadora) se liga ao DNA, a transcrição é estimulada. Alguns óperons são induzíveis; normalmente a transcrição está desligada e tem de ser ligada. Outros óperons são reprimíveis; a transcrição está ligada e tem de ser desligada.

Checagem dos conceitos 5

Em um óperon reprimível negativo, a proteína reguladora é sintetizada como:

a. Um ativador ativo.
b. Um ativador inativo.
c. Um repressor ativo.
d. Um repressor inativo.

Óperon *lac* de *E. coli*

Em 1961, François Jacob e Jacques Monod descreveram o "modelo do óperon" para o controle genético do metabolismo da lactose na *E. coli.* Esse trabalho e a pesquisa posterior sobre a genética do metabolismo da lactose estabeleceram o óperon como

a unidade básica do controle de transcrição nas bactérias. Apesar do fato de, naquele momento, não existirem métodos para determinar as sequências de nucleotídios, Jacob e Monod deduziram *geneticamente* a estrutura do óperon ao analisar as interações das mutações que interferiam com a regulação normal do metabolismo da lactose. Examinaremos os efeitos de algumas dessas mutações após observarmos como o óperon *lac* regula o metabolismo da lactose.

Metabolismo da lactose. A lactose é um importante carboidrato encontrado no leite; ela pode ser metabolizada pela bactéria *E. coli*, encontrada no intestino dos mamíferos. A lactose não se difunde com facilidade pela membrana celular da *E. coli*, e deve ser transportada ativamente para a célula por meio da proteína permease (**Figura 16.7**). Para utilizar a lactose como fonte de energia, a *E. coli* primeiro deve quebrá-la em glicose e galactose, uma reação catalisada pela enzima *β*-galactosidase. Essa enzima também converte lactose em alolactose, um composto que participa da regulação do metabolismo desse açúcar. Uma terceira enzima, tiogalactosídio transacetilase, também é produzida pelo óperon *lac*, mas sua função no metabolismo da lactose não está clara ainda. Uma possível função é a desintoxicação, evitando o acúmulo de tiogalactosídios que são transportados na célula junto com a lactose pela lactose permease.

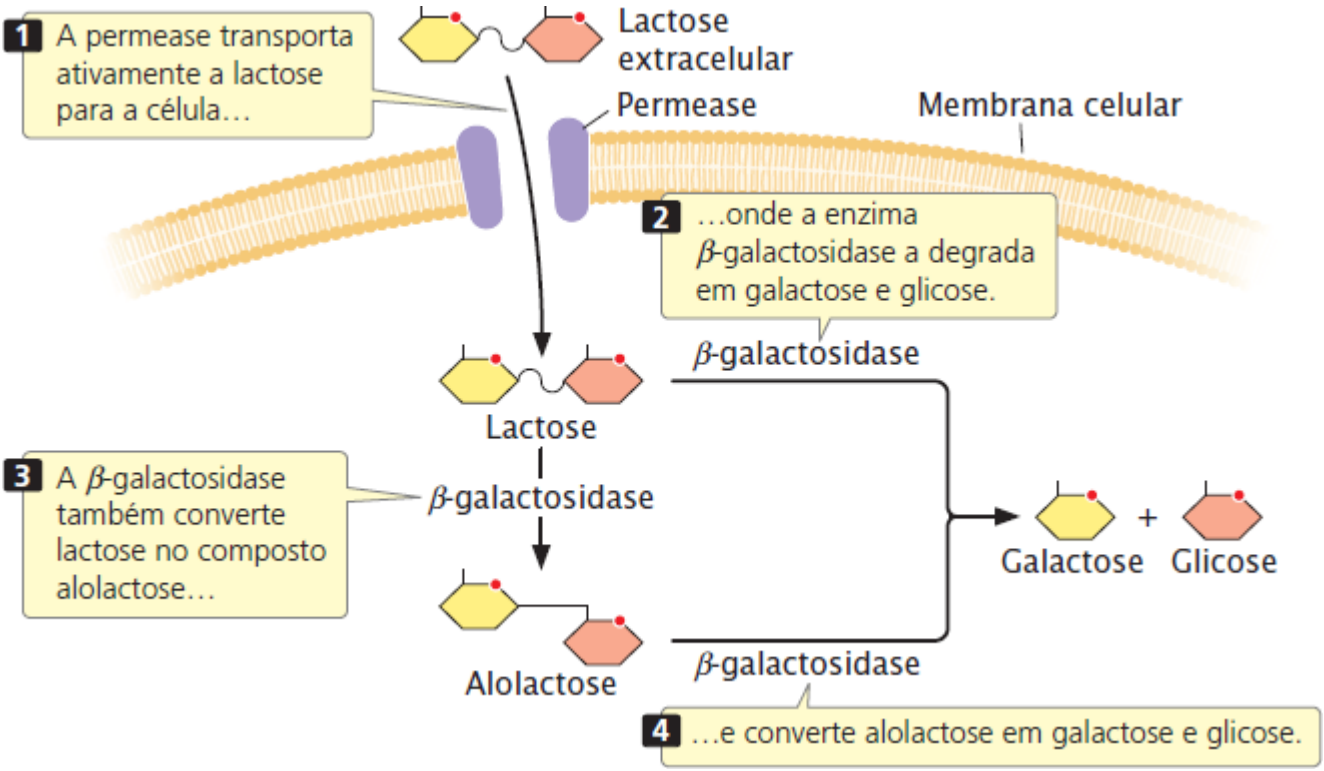


Figura 16.7 A lactose, um importante carboidrato encontrado no leite, é composta por dois açúcares de seis carbonos ligados entre si.

Regulação do óperon lac. O óperon *lac* é um exemplo de óperon induzível negativo. As enzimas *β*-galactosidase, permease e transacetilase são codificadas por genes estruturais adjacentes no óperon *lac* da *E. coli* (**Figura 16.8 A**) e têm um promotor em comum (*lacP* na Figura 16.8 A). A *β*-galactosidase é codificada pelo gene *lacZ*, a permease pelo gene *lacY*, e a transacetilase pelo gene *lacA*. Quando o meio no qual a *E. coli* cresce não tem lactose, são produzidas algumas moléculas de cada proteína. Se a lactose for adicionada ao meio e a glicose estiver ausente, a taxa de síntese de todas as três proteínas aumenta simultaneamente cerca de mil vezes dentro de 2 a 3 min. Essa explosão na síntese de proteínas é o resultado da transcrição de *lacZ*, *lacY* e *lacA* e exemplifica a **indução coordenada**, a síntese simultânea de várias proteínas, estimulada por uma molécula específica, o indutor (**Figura 16.8 B**).

Embora a lactose pareça ser o indutor aqui, a alolactose é de fato responsável pela indução. O gene regulador, *lacI*, está *upstream* do *lacP*, e tem seu próprio promotor (P_1). O gene *lacI* é transcrito em um mRNA curto que é traduzido em um repressor. O repressor é composto por quatro polipeptídios idênticos e tem dois tipos de sítios de ligação; um tipo de sítio se liga à alolactose, e o outro se liga ao DNA. Na ausência de lactose (e, portanto, alolactose), o repressor se liga ao sítio *lacO* do operador *lac* (ver Figura 16.8 A). Jacob e Monod mapearam o operador em uma posição adjacente ao gene do *lacZ*; a sequência de nucleotídios mais recente demonstrou que o operador, na verdade, se sobrepõe à extremidade 3′ do promotor e à extremidade 5′ do *lacZ* (**Figura 16.9**).

A RNA polimerase se liga ao promotor e se desloca na molécula de DNA, transcrevendo os genes estruturais. Quando o repressor está ligado ao operador, a ligação da RNA polimerase é bloqueada, impedindo a transcrição. Quando há lactose, parte

dela é convertida em alolactose, que se liga ao repressor e faz com que o repressor seja liberado do DNA. Havendo lactose, então, o repressor é inativado, a ligação da RNA polimerase não é mais bloqueada, ocorre a transcrição de *lacZ, lacY* e *lacA* e as proteínas *lac* são produzidas.

Você descobriu a falha na explicação dada para a indução das proteínas *lac*? Você se lembra de que a permease é necessária para transportar a lactose para a célula. Se o óperon *lac* for reprimido e não for produzida permease, como a lactose chega à célula para inativar o repressor e ligar a transcrição? Além disso, o indutor é a alolactose, que deve ser produzida a partir da lactose pela *β*-galactosidase. Se a produção da *β*-galactosidase é reprimida, como o metabolismo da lactose pode ser induzido?

A resposta é que a repressão nunca desliga *completamente* a transcrição do óperon *lac*. Mesmo com o repressor ativo ligado ao operador, existe um nível reduzido de transcrição e são sintetizadas algumas moléculas de *β*-galactosidase, permease e transacetilase. Quando a lactose aparece no meio, a permease existente transporta um pouco de lactose para a célula. As poucas moléculas de *β*-galactosidase encontradas convertem parte da lactose em alolactose, que induz a transcrição.

Vários compostos relacionados com a alolactose também podem se ligar ao repressor *lac* e induzir a transcrição do óperon *lac*. Um exemplo de indutor é o isopropiltiogalactosídio (IPTG). Embora o IPTG inative o repressor e possibilite a transcrição de *lacZ, lacY* e *lacA*, esse indutor não é metabolizado pela *β*-galactosidase; por esse motivo, o IPTG é usado em pesquisa para examinar os efeitos da indução, independente do metabolismo.

Conceitos

O óperon *lac* da *E. coli* controla a transcrição de três genes necessários no metabolismo da lactose: o gene *lacZ*, que codifica a *β*-galactosidase; o gene *lacY*, que codifica a permease e o gene *lacA*, que codifica a tiogalactosídio transacetilase. O óperon *lac* é induzível negativo: um gene regulador produz um repressor que se liga ao sítio do operador e evita a transcrição dos genes estruturais. A alolactose inativa o repressor e torna possível a transcrição do óperon *lac*.

Checagem dos conceitos 6

Quando há alolactose, o repressor *lac:*

a. Liga-se ao operador.
b. Liga-se ao promotor.
c. Não pode se ligar ao operador.
d. Liga-se ao gene regulador.

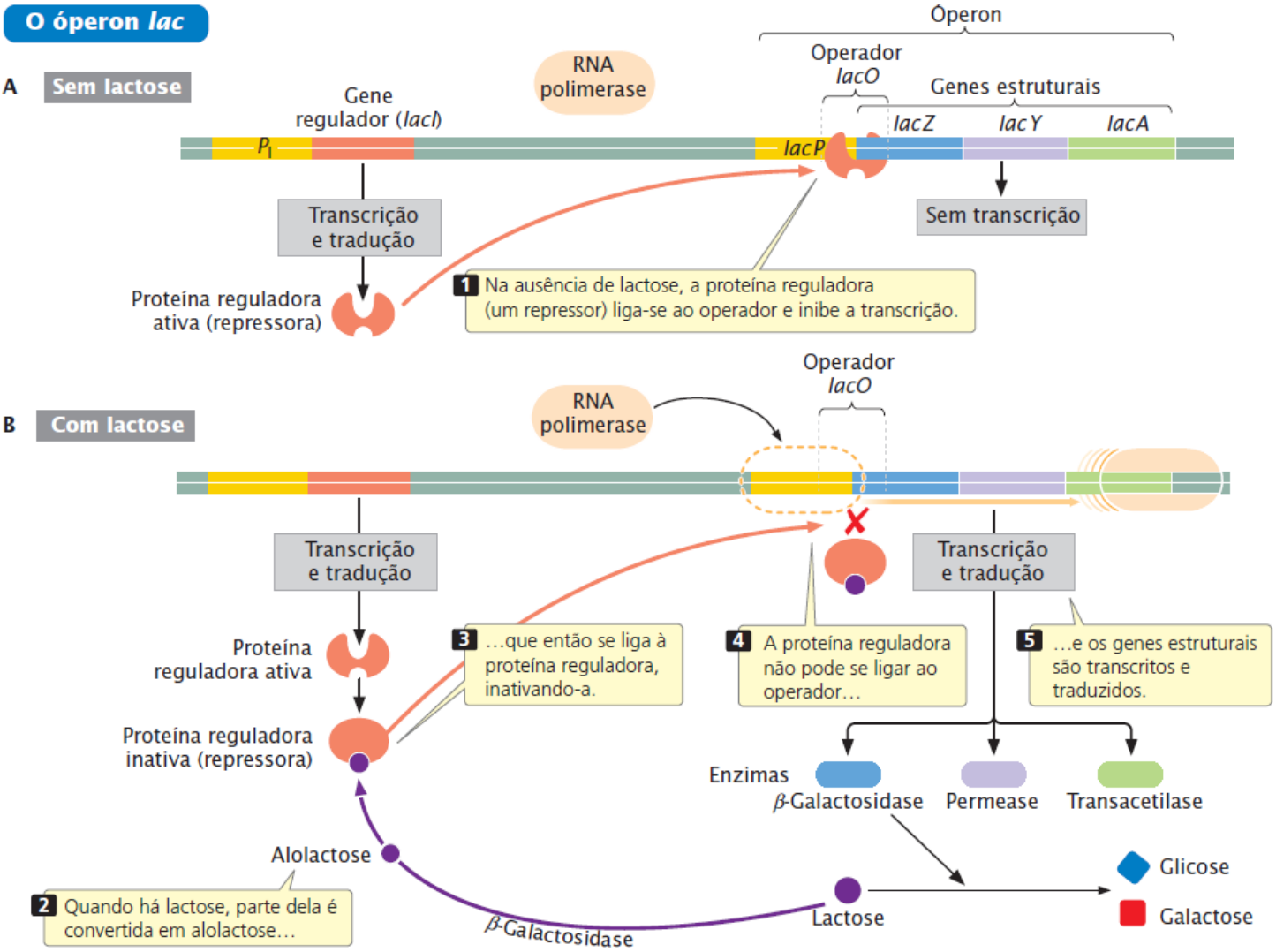


Figura 16.8 **O óperon *lac* regula o metabolismo da lactose.**

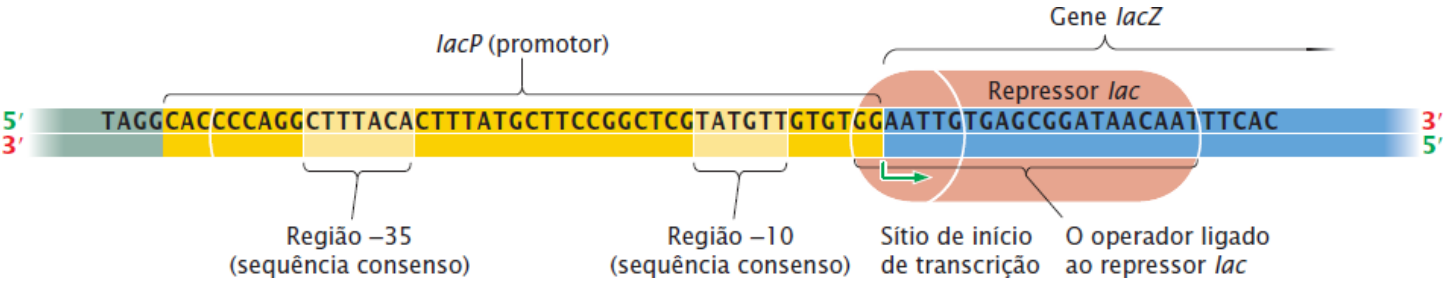


Figura 16.9 **No óperon *lac*, o operador se sobrepõe ao promotor e à extremidade 5′ do primeiro gene estrutural.**

Mutações *lac*

Jacob e Monod desenvolveram a estrutura e a função do óperon *lac* ao analisar mutações que afetavam o metabolismo da lactose. Para ajudar a definir os papéis dos diferentes componentes do óperon, eles usaram cepas **diploides parciais** de *E. coli*. As células dessas cepas têm duas moléculas de DNA diferentes: o cromossomo bacteriano completo e um pedaço extra de DNA. Jacob e Monod criaram essas cepas ao fazer com que ocorresse a conjugação entre duas bactérias (ver *Capítulo 9*). Na conjugação, um pequeno pedaço circular de DNA (o plasmídio F, ver *Capítulo 9*) é transferido de uma bactéria para outra. O plasmídio F usado por Jacob e Monod tinha um óperon *lac* de modo que a bactéria receptora se tornava parcialmente diploide, com duas cópias do óperon *lac*. Ao usar diferentes combinações de mutações no DNA bacteriano e do plasmídio, Jacob e Monod determinaram que algumas partes do óperon *lac* têm ação cis (capaz de controlar a expressão dos genes apenas quando no mesmo pedaço de DNA), enquanto outras partes têm ação trans (capaz de controlar a expressão dos genes em outras moléculas de DNA).

Mutações dos genes estruturais. Jacob e Monod descobriram primeiro algumas cepas mutantes que tinham perdido a capacidade de sintetizar β-galactosidase ou permease (eles não estudaram em detalhes os efeitos das mutações na enzima

transcetilase, e então a transcetilase não será estudada aqui). As mutações nas cepas mutantes foram mapeadas nos genes estruturais *lacZ* ou *lacY* e alteraram as sequências de aminoácidos das proteínas codificadas pelos genes. Essas mutações afetavam de forma clara a *estrutura* das proteínas, mas não a regulação da sua síntese.

Usando os diploides parciais, Jacob e Monod estabeleceram que as mutações nos genes *lacZ* e *lacY* eram independentes e afetavam apenas o produto do gene no qual a mutação ocorria. Os diploides parciais com $lacZ^+$ $lacY^-$ no cromossomo bacteriano e $lacZ^-$ $lacY^+$ no plasmídio funcionavam normalmente, produzindo *β*-galactosidase e permease quando há lactose. (O genótipo de um diploide parcial é escrito ao separar os genes em cada molécula de DNA com uma barra inclinada: $lacZ^+$ $lacY^-/lacZ^-$ $lacY^+$.) Nesse diploide parcial, um único gene funcional da *β*-galactosidase ($lacZ^+$) é suficiente para produzir *β*-galactosidase; não faz diferença se o gene funcional da *β*-galactosidase está acoplado a um gene funcional ($lacY^+$) ou defeituoso ($lacY^-$) da permease. O mesmo é válido para o gene $lacY^+$.

Mutações no gene regulador. Jacob e Monod também isolaram mutações que afetavam a *regulação* da produção de proteína. As mutações no gene *lacI* afetam a produção de *β*-galactosidase e permease porque os genes para ambas as proteínas estão no mesmo óperon e são regulados de forma coordenada.

Algumas dessas mutações são constitutivas, fazendo com que as proteínas *lac* sejam produzidas de forma contínua, independente de haver ou não lactose. Tais mutações no gene regulador são designadas como $lacI^-$. A construção de diploides parciais demonstrou que o gene $lacI^+$ é dominante sobre o gene $lacI^-$; uma única cópia de $lacI^+$ (genótipo $lacI^+/lacI^-$) era suficiente para a regulação normal da produção da proteína. Além disso, o $lacI^+$ restaurou o controle normal para um óperon mesmo se ele estivesse localizado em uma diferente molécula de DNA, mostrando que o $lacI^+$ pode ter ação trans. Um diploide parcial com genótipo $lacI^+$ $lacZ^-/lacI^-$ $lacZ^+$ funcionou normalmente, sintetizando *β*-galactosidase apenas quando havia lactose (**Figura 16.10**). Nessa cepa, o gene $lacI^+$ no cromossomo bacteriano era funcional, mas o gene $lacZ^-$ era defeituoso; no plasmídio, o gene $lacI^-$ era defeituoso, mas o gene $lacZ^+$ era funcional. O fato de que o gene $lacI^+$ pudesse regular um gene $lacZ^+$ localizado em uma molécula de DNA diferente indicou a Jacob e Monod que o produto do gene $lacI^+$ era capaz de funcionar no plasmídio ou no cromossomo.

Algumas mutações do *lacI* isoladas por Jacob e Monod evitaram que ocorresse a transcrição mesmo havendo lactose. Essas mutações eram chamadas de super-repressoras ($lacI^s$), porque produziam repressores defeituosos que não poderiam ser inativados por um indutor. As mutações $lacI^s$ produziam um repressor com um sítio de ligação de indutor alterado, que impedia o indutor de se ligar ao repressor; consequentemente, o repressor sempre era capaz de se fixar ao sítio do operador e evitar a transcrição dos genes *lac*. As mutações super-repressoras eram dominantes sobre $lacI^+$; os diploides parciais com genótipo $lacI^s$ $lacZ^+/lacI^+$ $lacZ^+$ não sintetizavam *β*-galactosidase ou permease, independente de haver ou não lactose (**Figura 16.11**).

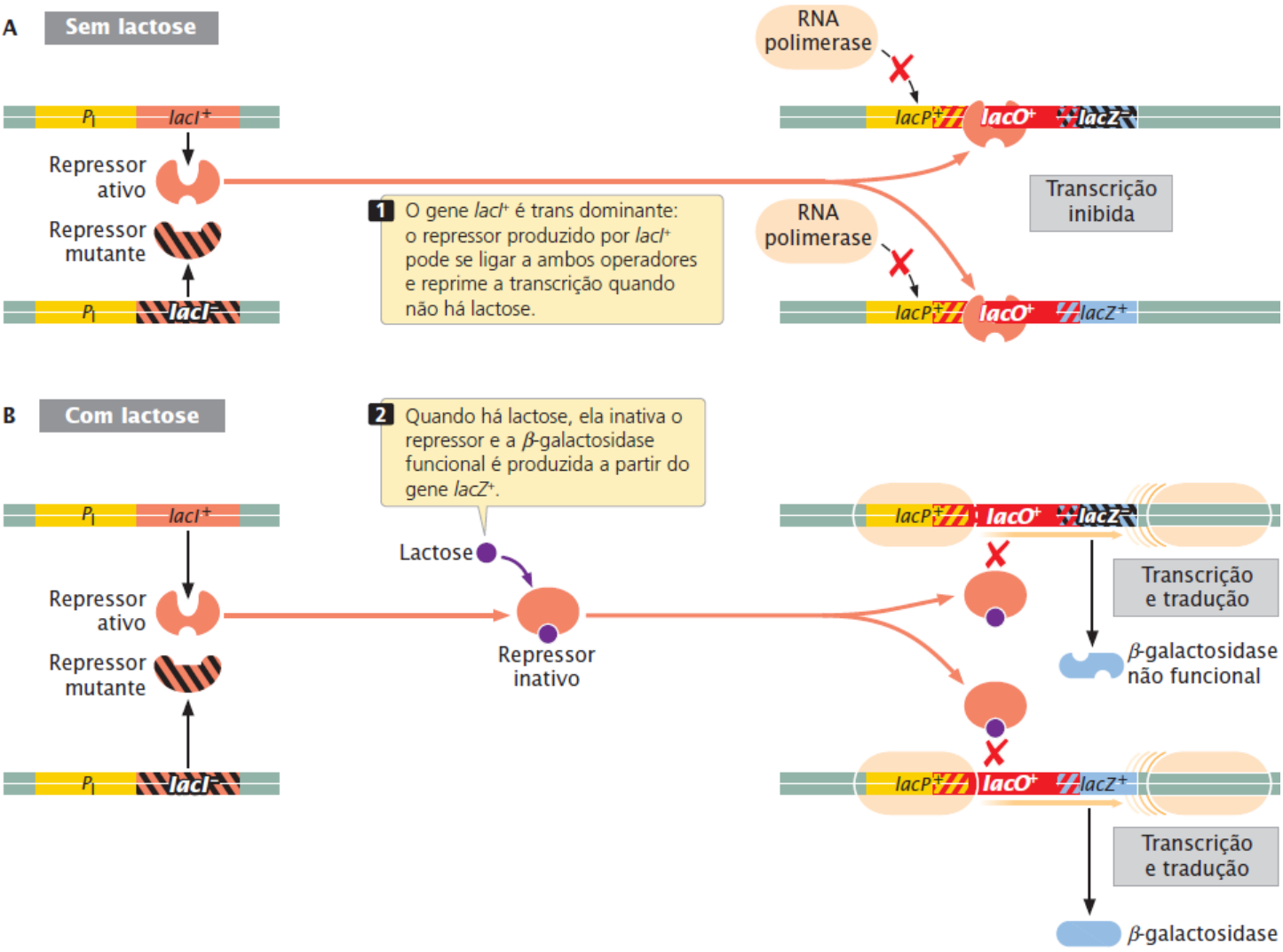


Figura 16.10 O diploide parcial *lacI*$^+$ *lacZ*$^-$/*lacI*$^-$ *lacZ*$^+$ produz *β*-galactosidase apenas quando há lactose, pois o gene *lacI* é trans dominante.

Mutações do operador. Jacob e Monod mapearam outra classe de mutantes constitutivos a um sítio adjacente ao *lacZ*. Essas mutações ocorriam no sítio do operador e eram chamadas de *lacO*^{c} ("O" representa operador, e "c", constitutivo). As mutações *lacO*^{c} alteravam a sequência de DNA no operador, de modo que a proteína repressora não era mais capaz de se ligar. Um diploide parcial com genótipo *lacI*$^+$ *lacO*^{c} *lacZ*$^+$/*lacI*$^+$ *lacO*$^+$ *lacZ*$^+$ exibia síntese constitutiva da *β*-galactosidase, indicando que o *lacO*^{c} é dominante sobre o *lacO*$^+$.

A análise de outros diploides parciais mostrou que o gene *lacO* é de ação cis, afetando apenas os genes na mesma molécula de DNA. Por exemplo, um diploide parcial com genótipo *lacI*$^+$ *lacO*$^+$ *lacZ*$^-$/*lacI*$^+$ *lacO*^{c} *lacZ*$^+$ era constitutivo, produzindo *β*-galactosidase havendo ou não lactose (**Figura 16.12 A**), mas um diploide parcial com genótipo *lacI*$^+$ *lacO*$^+$ *lacZ*$^+$/*lacI*$^+$ *lacO*^{c} *lacZ*$^-$ produziu *β*-galactosidase apenas quando havia lactose (**Figura 16.12 B**). O diploide parcial constitutivo (*lacI*$^+$ *lacO*$^+$ *lacZ*$^-$/*lacI*$^+$ *lacO*^{c} *lacZ*$^+$; ver Figura 16.12 A), a mutação *lacO*^{c} e o gene *lacZ*$^+$ funcional encontram-se na mesma molécula de DNA, mas, no *lacI*$^+$ *lacO*$^+$ *lacZ*$^+$/*lacI*$^+$ *lacO*^{c} *lacZ*$^-$ (ver Figura 16.12 B), a mutação *lacO*^{c} e o gene *lacZ*$^+$ funcional estão em moléculas diferentes. A mutação *lacO* afeta apenas os genes aos quais está fisicamente conectada, como em todas as mutações de operador. Elas evitam a ligação de uma proteína repressora ao operador e, portanto, tornam possível que a RNA polimerase transcreva genes na mesma molécula de DNA. Entretanto, eles não podem evitar que um repressor se ligue aos operadores normais em outras moléculas de DNA. Resolva o Problema 21

Mutações do promotor. As mutações que afetam o metabolismo da lactose também foram isoladas no sítio do promotor; essas mutações são chamadas de *lacP*$^-$ e interferem com a ligação da RNA polimerase ao promotor. Como tal ligação é essencial para a transcrição dos genes estruturais, as cepas de *E. coli* com mutações *lacP*$^-$ não produzem as proteínas *lac* quando há nem quando não há lactose. Como as mutações do operador, as mutações *lacP*$^-$ têm ação cis e afetam apenas os genes na mesma molécula de DNA. O diploide parcial *lacI*$^+$ *lacP*$^+$ *lacZ*$^+$/*lacI*$^+$ *lacP*$^-$ *lacZ*$^+$ exibe a síntese normal de *β*-galactosidase, enquanto o *lacI*$^+$ *lacP*$^-$ *lacZ*$^+$/*lacI*$^+$ *lacP*$^+$ *lacZ*$^-$ não produz essa enzima, havendo ou não lactose. O **Quadro 16.2** apresenta os diferentes tipos de mutações que ocorrem no óperon *lac*.

Problema desenvolvido

Para as cepas de *E. coli* com os genótipos *lac* a seguir, use um sinal de mais (+) para indicar a síntese de *β*-galactosidase e permease e um sinal de menos (−) para indicar sem síntese das proteínas havendo ou não lactose.

Genótipo da cepa

a. $lacI^+\ lacP^+\ lacO^+\ lacZ^+\ lacY^+$
b. $lacI^+\ lacP^+\ lacO^c\ lacZ^-\ lacY^+$
c. $lacI^+\ lacP^-\ lacO^+\ lacZ^+\ lacY^-$
d. $lacI^+\ lacP^+\ lacO^+\ lacZ^-\ lacY^-/lacI^-\ lacP^+\ lacO^+\ lacZ^+\ lacY^+$

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

Uma indicação se a *β*-galactosidase e a permease são produzidas ou não por cada genótipo quando há ou não lactose, inserindo um sinal de (+) ou de menos (−) para cada enzima e condição no quadro.

Quais informações são fornecidas para solucionar o problema?

O genótipo de cada cepa.

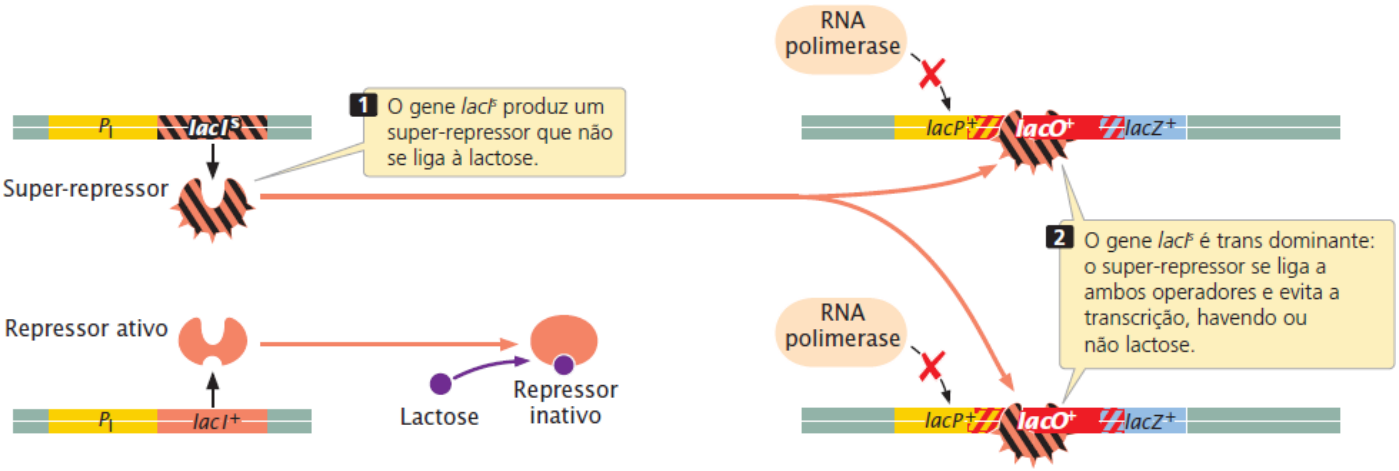


Figura 16.11 O diploide parcial *lacI^S lacZ^+/lacI^+ lacZ^+* não produz *β*-galactosidase havendo ou não lactose, pois o gene *lacI^S* codifica um super-repressor.

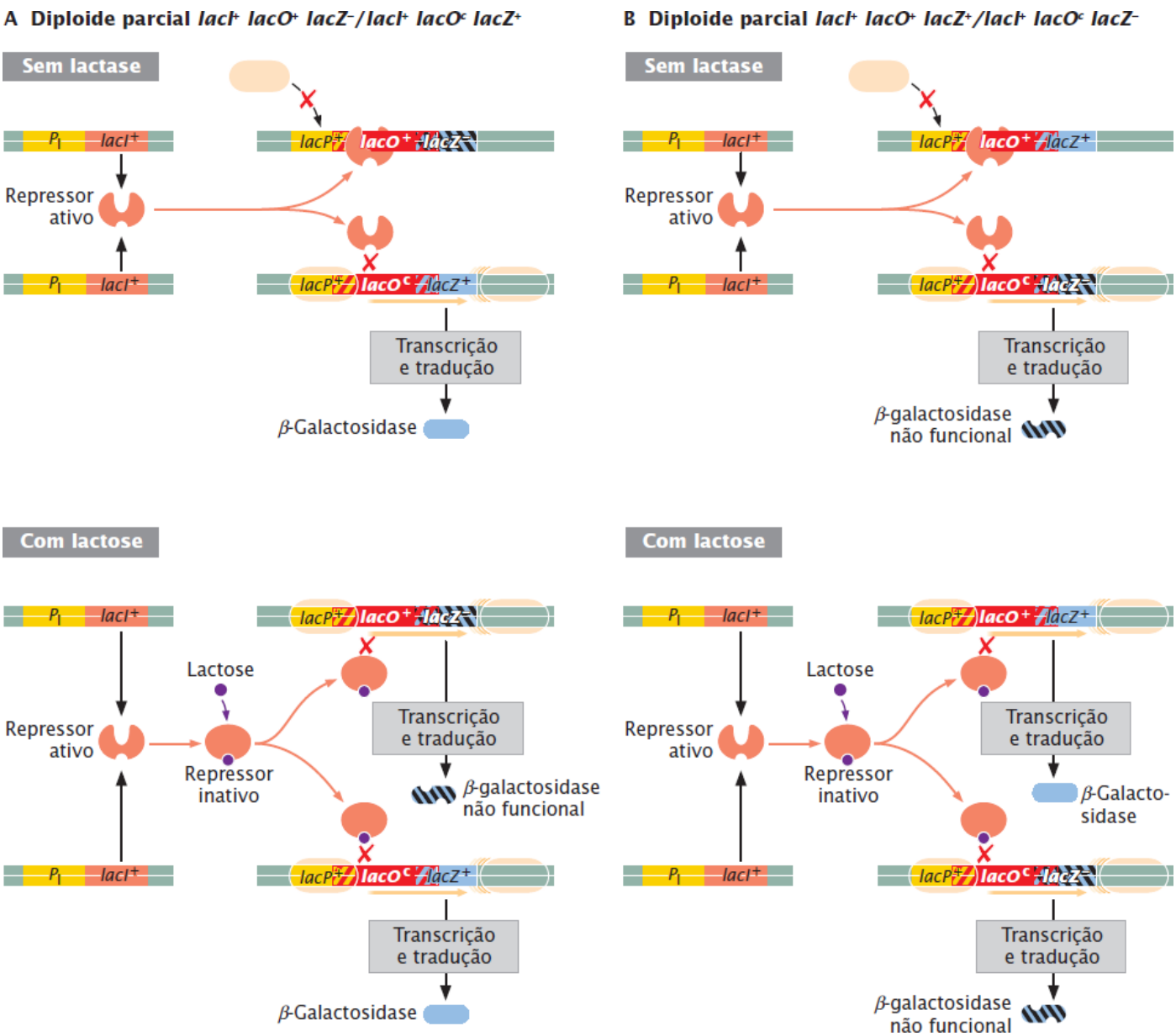


Figura 16.12 **As mutações no *lacO* são constitutivas e com ação cis. A.** O diploide parcial *lacI*$^+$ *lacO*$^+$ *lacZ*$^-$/*lacI*$^+$ *lacO*^{c} *lacZ*$^+$ é constitutivo, produzindo *β*-galactosidase, havendo ou não lactose. **B.** O diploide parcial *lacI*$^+$ *lacO*$^+$ *lacZ*$^+$/*lacI*$^+$ *lacO*^{c} *lacZ*$^-$ é induzível (produz *β*-galactosidase apenas quando há lactose), demonstrando que o gene *lacO* tem ação cis.

Quadro 16.2 Características das mutações do óperon *lac*.

Tipo	Localização	Cis/Trans	Efeito
Mutações do gene estrutural	*lacZ, lacY*	Afetam apenas *lacZ* e *lacY*	Alteram a sequência de aminoácidos da proteína codificada pelo gene no qual ocorre a mutação
Mutações do gene regulador	*lacI*	Trans	Afetam a transcrição dos genes estruturais
Mutações do operador	*lacO*	Cis	Afetam a transcrição dos genes estruturais
Mutações do promotor	*lacP*	Cis	Afetam a transcrição dos genes estruturais

Etapas para a solução

Genótipo da cepa	Sem lactose		Com lactose	
	β-galactosidase	**Permease**	**β-galactosidase**	**Permease**
a. *lacI*$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^+$	−	−	+	+
b. *lacI*$^+$ *lacP*$^+$ *lacO*^{c} *lacZ*$^-$ *lacY*$^+$	−	+	−	+
c. *lacI*$^+$ *lacP*$^-$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^-$	−	−	−	−
d. *lacI*$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^-$ *lacY*$^-$/*lacI*$^-$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^+$	−	−	+	+

a. Todos os genes têm sequências normais, então o óperon *lac* funciona normalmente: quando não há lactose, a proteína reguladora se liga ao operador e inibe a transcrição dos genes estruturais, e então a *β*-galactosidase e a permease não são produzidas. Quando há lactose, parte dela é convertida em alolactose, que se liga ao repressor e a inativa; o repressor não se liga ao operador, e então os genes estruturais são transcritos e a *β*-galactosidase e a permease são produzidas.
b. O gene *lacZ* estrutural é mutante, de modo que a *β*-galactosidase não será produzida sob nenhuma condição. O gene *lacO* tem uma mutação constitutiva, o que significa que o repressor é incapaz de se ligar ao *lacO*, e a transcrição ocorre em qualquer circunstância. Portanto, a permease será produzida, havendo ou não lactose.
c. Nesta cepa, o promotor é mutante, e então a RNA polimerase não pode se ligar e não ocorre a transcrição. Desta forma, a *β*-galactosidase e a permease não são produzidas sob nenhuma condição.
d. Esta cepa é um diploide parcial, composto por duas cópias do óperon *lac* – uma em um cromossomo bacteriano e outra em um plasmídio. O óperon *lac* representado na parte superior do genótipo tem mutações nos genes *lacZ* e *lacY*, não sendo capaz de codificar a *β*-galactosidase ou a permease em nenhuma circunstância. O óperon *lac* na parte inferior do genótipo tem um gene regulador defeituoso, mas o gene regulador normal no óperon superior produz um repressor difusível (ação *trans*) que se liga ao óperon inferior na ausência de lactose e inibe a transcrição. Portanto, não é produzida *β*-galactosidase nem permease quando não há lactose. Havendo lactose, o repressor não pode se ligar ao operador; então, o óperon inferior é transcrito e a *β*-galactosidase e a permease são produzidas.

Agora preveja o desfecho das diferentes mutações *lac* ao resolver o **Problema 19** no final do capítulo.

Controle positivo e repressão catabólica

E. coli e muitas outras bactérias metabolizam a glicose preferencialmente quando há lactose e outros açúcares. Isto é possível porque a glicose entra na glicólise sem modificação e, portanto, requer menos energia para metabolizar do que outros açúcares. Quando a glicose está disponível, os genes que participam do metabolismo dos outros açúcares são reprimidos, em um fenômeno conhecido como **repressão catabólica**. Por exemplo, a transcrição eficiente do óperon *lac* ocorre apenas se houver lactose e se a glicose estiver ausente. No entanto, como a expressão do óperon *lac* é influenciada pela glicose? O que produz a repressão catabólica?

A repressão catabólica surge a partir do controle positivo em resposta à glicose. (Essa regulação é uma adição para o controle negativo produzido pelo repressor ligado ao sítio do operador do óperon *lac* quando a lactose estiver ausente.) O controle positivo é obtido por meio da ligação de uma proteína dimérica chamada **proteína catabólica ativadora** (CAP) a um sítio que tem cerca de 22 nucleotídios de comprimento e está localizado dentro ou discretamente *upstream* do promotor dos genes *lac* (**Figura 16.13**). A RNA polimerase não se liga de forma eficiente a muitos promotores, a menos que a CAP esteja ligada primeiro no DNA. Antes que a CAP possa se ligar ao DNA, ela deve formar um complexo com um nucleotídio modificado chamado **monofosfato de adenosina-3′,5′-cíclico** (AMP cíclico ou cAMP), importante nos processos de sinalização das células bacterianas e eucarióticas. Na *E. coli*, a concentração de cAMP é regulada de modo que sua concentração é inversamente proporcional ao nível de glicose disponível. Uma alta concentração de glicose dentro da célula reduz tanto a quantidade de cAMP que o complexo cAMP-CAP fica disponível para se ligar ao DNA. Consequentemente, a RNA polimerase tem pouca afinidade pelo promotor *lac*, e ocorre pouca transcrição do óperon *lac*. Baixas concentrações da glicose estimulam níveis elevados de cAMP, resultando em aumento da ligação cAMP-CAP ao DNA. Esse aumento potencializa a ligação da RNA polimerase ao promotor e aumenta a transcrição dos genes *lac* em aproximadamente 50 vezes.

A proteína ativadora catabólica exerce controle positivo em mais de 20 óperons da *E. coli.* A resposta à CAP varia entre esses promotores; alguns óperons são ativados por níveis reduzidos de CAP, enquanto outros exigem níveis elevados. A CAP tem um motivo de ligação ao DNA hélice-giro-hélice e quando ela se liga ao sítio CAP no DNA, fazendo com que a hélice do DNA se dobre (**Figura 16.14**). A hélice dobrada permite que a CAP se ligue à RNA polimerase no promotor e no início da transcrição.

Resolva o Problema 14

Conceitos

Apesar do seu nome, a repressão catabólica é um tipo de controle positivo no óperon *lac*. A proteína catabólica ativadora (CAP), complexada com cAMP, liga-se ao sítio próximo do promotor e estimula a ligação da RNA polimerase. Níveis celulares de cAMP são controlados pela glicose; um baixo nível de glicose aumenta a abundância de cAMP e potencializa a transcrição dos genes estruturais do *lac*.

Checagem dos conceitos 7

Qual é o efeito dos níveis elevados de glicose no óperon *lac*?

a. A transcrição é estimulada.
b. Ocorre pouca transcrição.
c. A transcrição não é afetada.
d. A transcrição pode ser estimulada ou inibida, dependendo dos níveis de lactose.

Óperon *trp* de *E. coli*

O óperon *lac* já discutido é um óperon induzível, no qual a transcrição não ocorre normalmente e deve ser ligada. Alguns óperons são reprimíveis; a transcrição nesses óperons está ligada e deve ser reprimida. O óperon do triptofano (*trp*) na *E. coli*, que controla a biossíntese do aminoácido triptofano, é um exemplo de óperon reprimível negativo.

O óperon *trp* tem cinco genes estruturais (*trpE, trpD, trpC, trpB* e *trpA*) que produzem os componentes das três enzimas (duas das enzimas são compostas por duas cadeias de polipeptídios). Essas enzimas convertem corismato em triptofano (**Figura 16.15**). O primeiro gene estrutural, *trpE*, tem uma região 5′ não traduzida (5′ UTR) longa que é transcrita, mas não codifica nenhuma dessas enzimas. Pelo contrário, essa 5′ UTR participa de outro mecanismo regulatório, discutido na Seção a seguir. O promotor *trp* está *upstream* da 5′ UTR. Quando os níveis de triptofano estão baixos, a RNA polimerase se liga ao promotor e transcreve os cinco genes estruturais em um único mRNA, que é então traduzido em enzimas que convertem o corismato em triptofano.

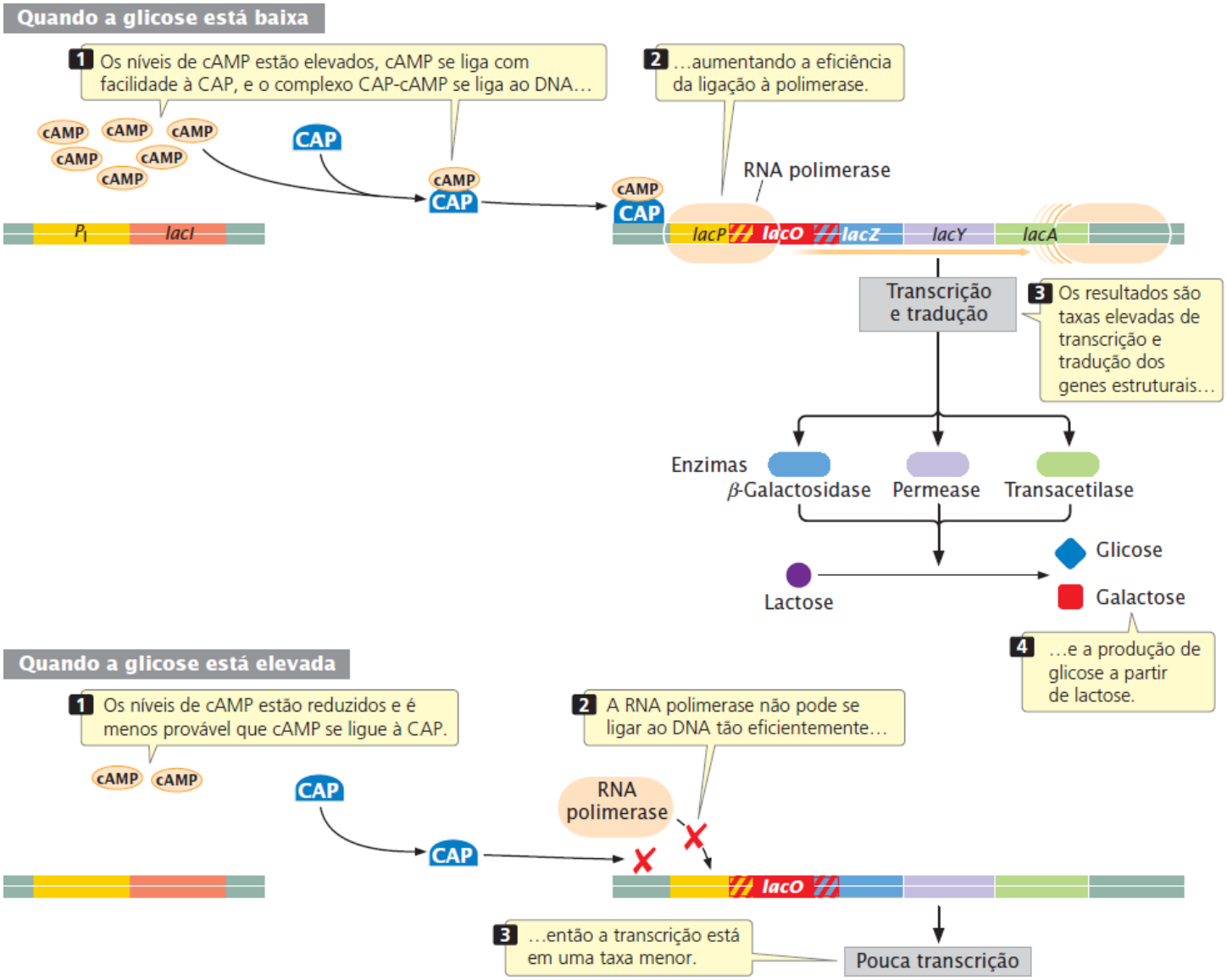


Figura 16.13 A proteína catabólica ativadora (CAP) se liga ao promotor do óperon *lac* e estimula a transcrição. A CAP deve se complexar com o monofosfato de adenosina-3′,5′-cíclico (cAMP) antes de se ligar ao promotor do óperon *lac*. A ligação do cAMP-CAP ao promotor ativa a transcrição ao facilitar a ligação da RNA polimerase. Os níveis de cAMP estão inversamente relacionados com a glicose: glicose baixa estimula cAMP elevado; glicose elevada estimula cAMP baixo.

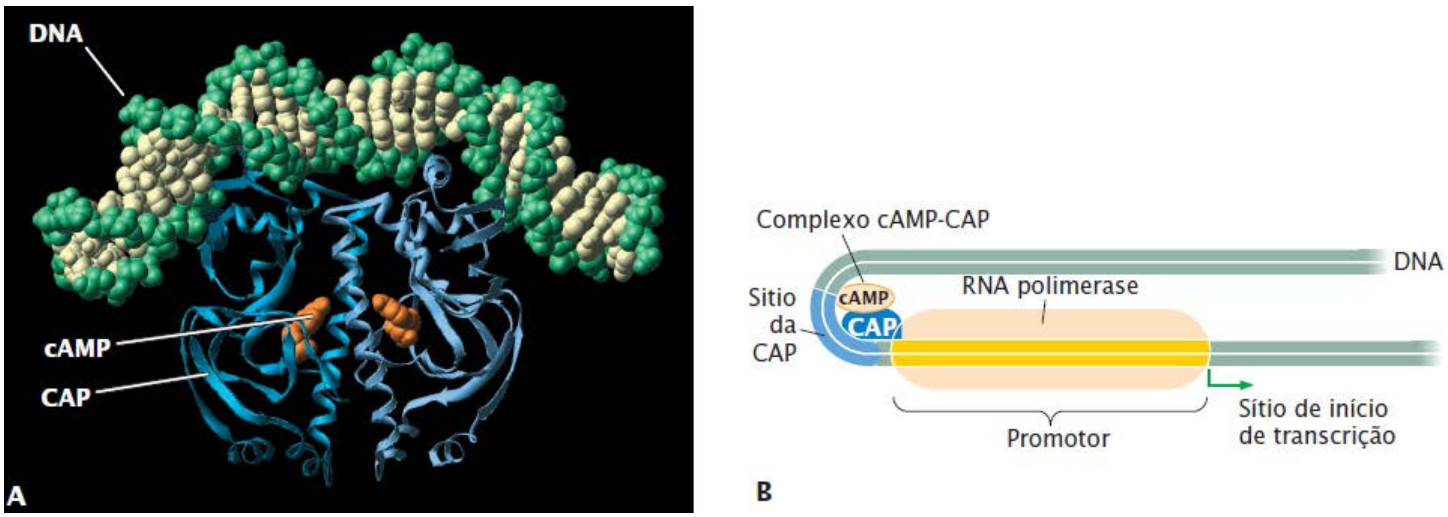


Figura 16.14 A ligação do complexo cAMP-CAP ao DNA produz uma dobra acentuada no DNA que ativa a transcrição.

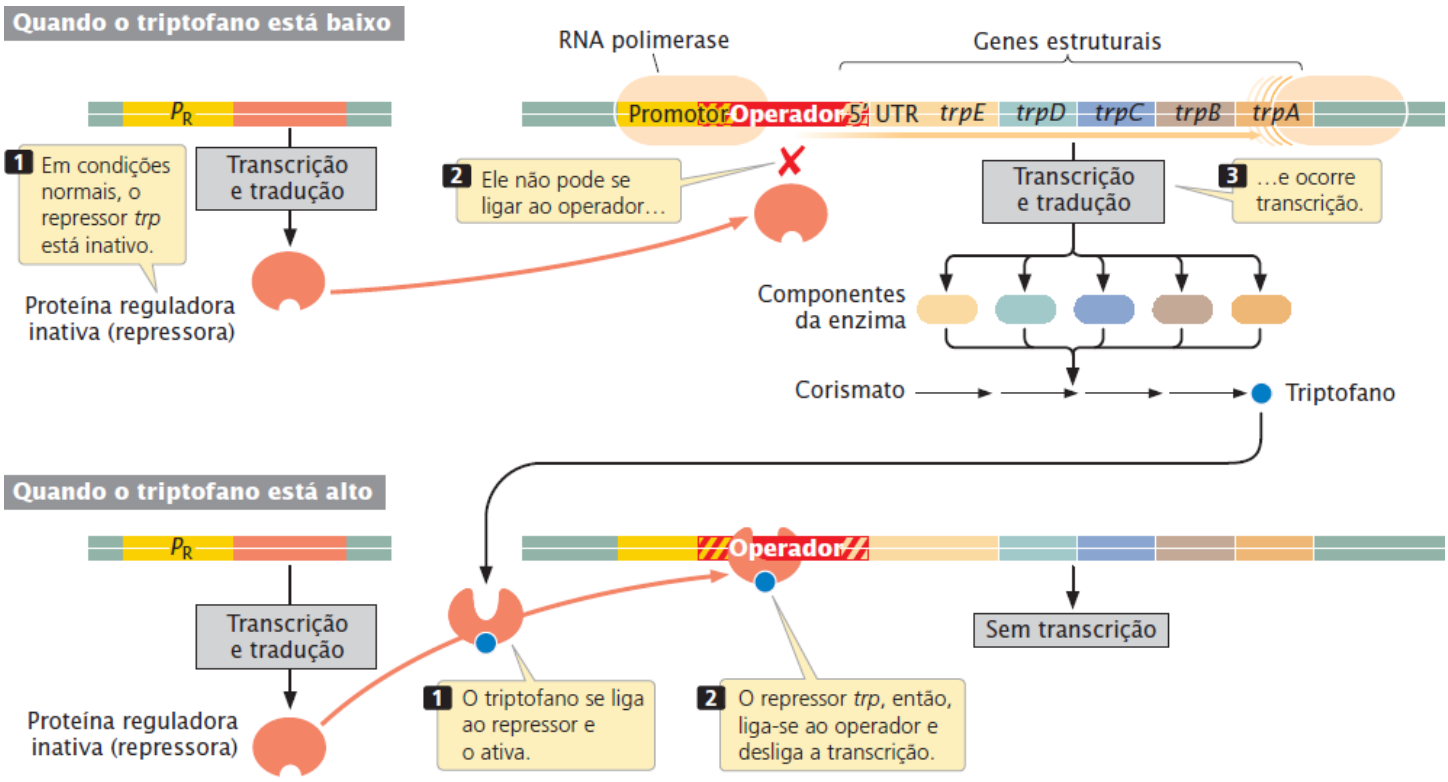


Figura 16.15 **O óperon *trp* controla a biossíntese do aminoácido triptofano na *E. coli*.**

A uma certa distância do óperon *trp* está um gene regulador, o *trpR*, que codifica um repressor que sozinho não pode se ligar ao DNA (ver Figura 16.15). Como o repressor *lac*, o repressor do triptofano tem dois sítios de ligação, um que se liga ao DNA no sítio do operador e outro que se liga ao triptofano (o ativador). A ligação com o triptofano leva a uma mudança conformacional no repressor, que o torna capaz de se ligar ao DNA no sítio do operador, que se sobrepõe ao promotor. Quando o operador está ocupado pelo repressor do triptofano, a RNA polimerase não pode se ligar ao promotor e os genes estruturais não podem ser transcritos. Assim, quando os níveis celulares de triptofano estão baixos, ocorre a transcrição do óperon *trp* e mais triptofano é sintetizado; quando os níveis celulares de triptofano estão elevados, a transcrição do óperon *trp* é inibida e não ocorre mais transcrição desse aminoácido.

Acentuadores bacterianos

Outro tipo de sequência regulatória que afeta a transcrição é um acentuador, um elemento do DNA que afeta a transcrição, mas, ao contrário dos promotores, é encontrado a certa distância do gene (ver *Capítulo 17*). Os acentuadores foram descritos originalmente nos eucariotos, mas a pesquisa indica que também são encontrados em bactérias e archaea.

Como os acentuadores nos eucariotos, os acentuadores bacterianos têm sítios de ligação para proteínas que aumentam a taxa de transcrição a partir dos promotores que estão distantes do gene. Isto é possível ao fazer com que o DNA entre o promotor e acentuador crie uma alça, de modo que o fator de transcrição no acentuador interaja diretamente com a RNA polimerase no promotor. Os acentuadores também são independentes de posição, o que significa que eles podem ser deslocados sem afetar sua capacidade de potencializar a transcrição. A maioria dos acentuadores bacterianos é encontrada *upstream* dos genes que utilizam um tipo especial de fator sigma (ver *Capítulo 13*), conhecido como sigma 54 (σ^{54}). Os acentuadores serão discutidos com mais detalhes no Capítulo 17.

Conceitos

O óperon *trp* é um óperon reprimível negativo que controla a biossíntese do triptofano. Em um óperon reprimível, a transcrição está normalmente ligada e tem de ser reprimida: esse processo é feito por meio da ligação do triptofano ao repressor, que torna o repressor ativo. O repressor ativo se liga ao operador e evita que a RNA polimerase transcreva os genes estruturais. Os acentuadores bacterianos aumentam a taxa de transcrição em genes distantes do acentuador.

Checagem dos conceitos 8

No óperon *trp*, o que acontece ao repressor *trp* na ausência de triptofano?

a. Ele se liga ao operador e reprime a transcrição.

b. Ele não pode se ligar ao operador e ocorre a transcrição.
c. Ele se liga ao gene regulador e reprime a transcrição.
d. Ele não pode se ligar ao regulador e ocorre a transcrição.

16.3 Alguns óperons regulam a transcrição por meio da atenuação, o término prematuro da transcrição

Agora veremos as diferentes formas pelas quais uma célula regula o início da transcrição em um óperon. Alguns óperons têm um nível de controle adicional que afeta a *continuação* da transcrição em vez do seu início. Na **atenuação**, a transcrição começa no sítio de início, mas o término é precoce, antes que a RNA polimerase alcance os genes estruturais. A atenuação ocorre em vários óperons que codificam as enzimas participantes da biossíntese dos aminoácidos.

Atenuação no óperon *trp* da *E. coli*

Podemos compreender o processo de atenuação mais facilmente ao examinar um dos exemplos mais estudados, que é encontrado no óperon *trp* da *E. coli*. O óperon *trp* é incomum por ser regulado tanto pela repressão quanto pela atenuação. A maioria dos óperons é regulada por um desses mecanismos, mas não por ambos.

A atenuação surgiu pela primeira vez quando Charles Yanofsky e seus colegas fizeram várias observações no início dos anos 1970 que indicavam que a repressão no sítio operador não é o único método de regulação no óperon *trp*. Eles isolaram vários mutantes que exibiam níveis elevados de transcrição, e o controle no sítio do operador não foi afetado, sugerindo que algum mecanismo além da repressão no sítio do operador estava controlando a transcrição. Além disso, eles observaram que dois mRNAs de diferentes tamanhos foram transcritos a partir do óperon *trp*: um longo mRNA com sequências para os genes estruturais e um mRNA muito mais curto de apenas 140 nucleotídios. Essas observações fizeram Yanofsky propor que o mecanismo responsável pelo término prematuro da transcrição também regula a transcrição do óperon *trp*.

Um exame detalhado do óperon *trp* revela uma região de 162 nucleotídios que corresponde à 5′ UTR longa do mRNA (já mencionada) transcrita a partir do óperon *trp* (**Figura 16.16 A**). A 5′ UTR (também chamada de líder) tem quatro regiões: a região 1 é complementar à região 2, a região 2 é complementar à 3, e a região 3 é complementar à região 4 (**Figura 16.16 B**). Essas complementaridades fazem com que a 5′ UTR se dobre em duas estruturas secundárias diferentes (**Figura 16.16 C**). Uma dessas estruturas secundárias é responsável pela atenuação.

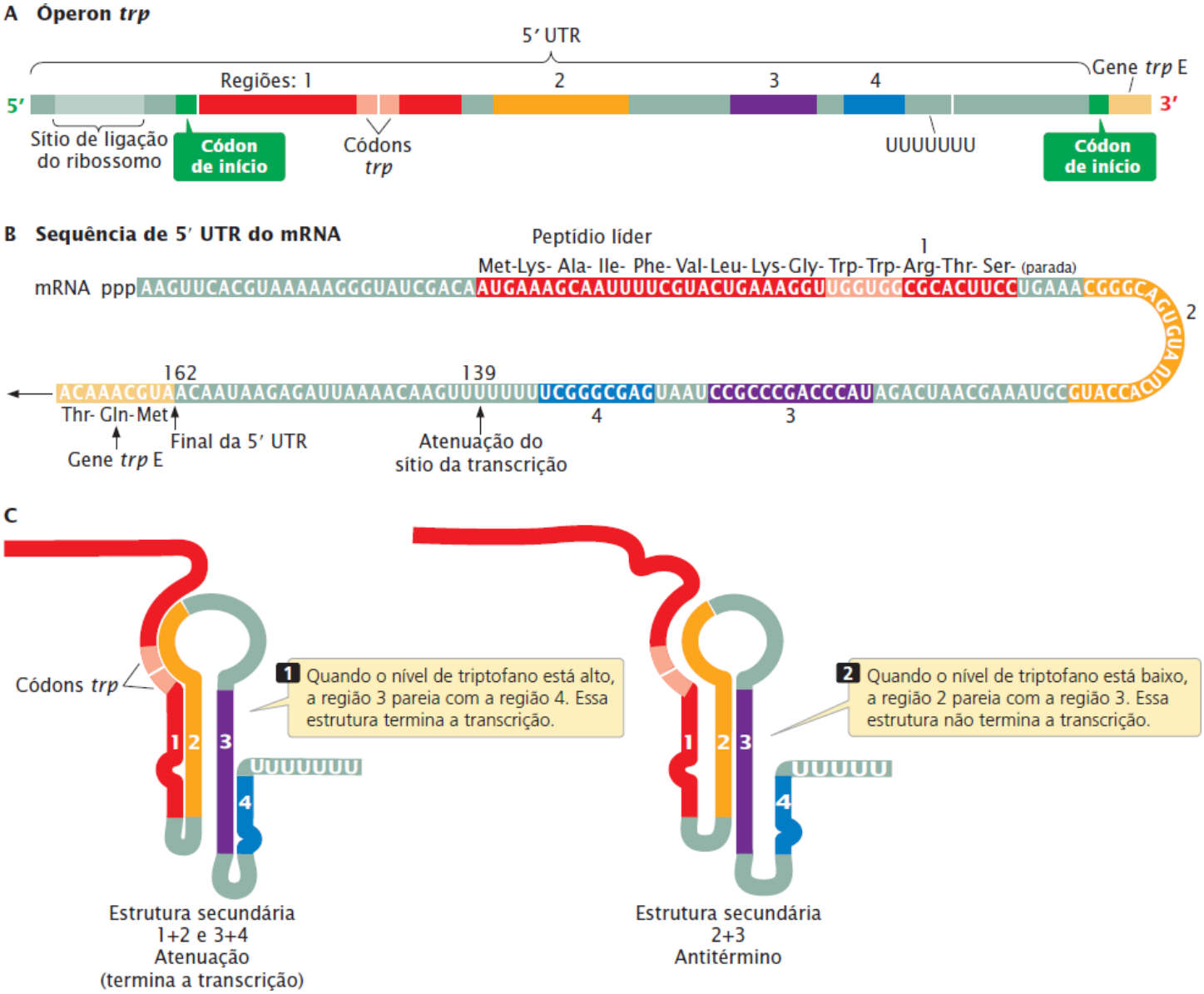


Figura 16.16 Duas estruturas secundárias diferentes podem ser formadas pela 5′ UTR do mRNA transcrito do óperon *trp*.

Uma das estruturas secundárias tem um grampo (*hairpin*) produzido pelo pareamento de bases das regiões 1 e 2 e outro grampo produzido pelo pareamento de bases das regiões 3 e 4. Observe que uma fita de nucleotídios uracila acompanha a estrutura em grampo 3+4. Não coincidentemente, a estrutura do terminador intrínseco bacteriano (ver *Capítulo 13*) inclui um grampo seguido de uma fita de nucleotídios uracila; essa estrutura secundária na 5′ UTR do óperon *trp* é na verdade um terminador, e é chamado de **atenuador**. O atenuador se forma quando os níveis celulares de triptofano estão elevados, terminando a transcrição antes que os genes estruturais *trp* possam ser transcritos.

Quando os níveis celulares de triptofano estão baixos, entretanto, uma estrutura secundária alternativa da 5′ UTR é produzida pelo paramento de bases nas regiões 2 e 3 (ver Figura 16.16 B). Esse pareamento de bases também produz um grampo, mas não acompanhado por uma fita de nucleotídios uracila; então, essa estrutura *não* funciona como terminador. A RNA polimerase continua após a 5′ UTR na seção de codificação dos genes estruturais, e são produzidas as enzimas que sintetizam o triptofano. Como ela impede o término da transcrição, a estrutura 2+3 é chamada de **antiterminadora**.

Para resumir, a 5′ UTR do óperon *trp* pode se dobrar em uma das duas estruturas. Quando o nível de triptofano está elevado, a estrutura 3+4 se forma, a transcrição é terminada dentro da 5′ UTR e não é sintetizado mais triptofano. Quando o nível de triptofano está baixo, a estrutura 2+3 se forma, a transcrição continua por meio dos genes estruturais, e o triptofano é sintetizado. A pergunta crítica é: por que a estrutura 3+4 surge quando o nível de triptofano na célula está elevado, enquanto a estrutura 2+3 surge quando o nível está baixo?

Precisamos observar com mais detalhes a sequência de nucleotídios da 5′ UTR para responder a essa pergunta. Na extremidade 5′, *upstream* da região 1, existe um sítio de ligação do ribossomo (ver Figura 16.16 A). A região 1 codifica uma pequena proteína. Na sequência codificadora para essa proteína existem dois códons UGG, que especificam o aminoácido triptofano; então, o triptofano é necessário para a tradução dessa sequência 5′ UTR. A pequena proteína codificada pela 5′ UTR não foi isolada e acredita-se que seja instável; sua função aparente é apenas controlar a atenuação. Embora tenhamos afirmado

no Capítulo 14 que a 5′ UTR não é traduzida em uma proteína, a 5′ UTR dos óperons sujeitos à atenuação é uma exceção a essa regra. A cronologia exata e a interação da transcrição e tradução na 5′ UTR determinam se ocorre atenuação.

Transcrição quando os níveis de triptofano estão baixos. Consideremos primeiro o que acontece quando os níveis intracelulares de triptofano estão baixos. Lembre-se de que, nas células procarióticas, a transcrição e a tradução estão acopladas: enquanto ocorre a transcrição na extremidade 3′ do mRNA, a tradução é iniciada na extremidade 5′. A RNA polimerase começa a transcrever o DNA, produzindo a região 1 da 5′ UTR (**Figura 16.17 A**). Acompanhando de perto a RNA polimerase, um ribossomo se liga a 5′ UTR e começa a tradução da região codificadora. Enquanto isso, a RNA polimerase está transcrevendo a região 2 (**Figura 16.17 B**). A região 2 é complementar à região 1, mas, como o ribossomo está traduzindo a região 1, os nucleotídios nas regiões 1 e 2 não podem parear.

A RNA polimerase começa a transcrever a região 3, e o ribossomo alcança os códons UGG do triptofano na região 1. Quando ela alcança os códons do triptofano, o ribossomo para (**Figura 16.17 C**), porque o nível de triptofano está baixo e os tRNAs carreados com triptofano estão escassos ou até indisponíveis. O ribossomo se fixa nos códons do triptofano, esperando a chegada de um tRNA carreado com o aminoácido. A parada do ribossomo não inibe a transcrição; a RNA polimerase continua a se mover pelo DNA, e a transcrição ultrapassa a tradução.

Como o ribossomo está parado nos códons do triptofano na região 1, a região 2 está livre para parear as bases com a região 3, formando a estrutura em forma de grampo 2+3 (**Figura 16.17 D**). Esse grampo não leva ao término, então a transcrição continua. Como a região 3 já está pareada com a região 2, o grampo de cabelo 3+4 (o atenuador) nunca se forma, e então não ocorre a atenuação e a transcrição continua. A RNA polimerase continua ao longo do DNA, passa a 5′ UTR ao transcrever todos os genes estruturais em mRNA, que é traduzido nas enzimas codificadas pelo óperon *trp*. Essas enzimas, então, sintetizam mais triptofano.

Transcrição quando os níveis de triptofano estão elevados. Agora veremos o que acontece quando os níveis intracelulares de triptofano estão elevados. Mais uma vez, a RNA polimerase começa a transcrever o DNA, produzindo a região 1 da 5′ UTR (**Figura 16.17 E**). Acompanhando de perto a RNA polimerase, um ribossomo se liga à 5′ UTR e começa a tradução da região codificadora (**Figura 16.17 F**). Quando o ribossomo alcança os dois códons UGG do triptofano, ele não fica lento nem para, porque o triptofano está abundante e os tRNAs carreados com triptofano estão disponíveis (**Figura 16.17 G**). É importante observar esse ponto: como o triptofano está abundante, a tradução pode acompanhar a transcrição.

À medida que ele se move pela região 1, o ribossomo cobre parcialmente a região 2 (**Figura 16.17 H**); enquanto isso, a RNA polimerase completa a transcrição da região 3. Embora as regiões 2 e 3 sejam complementares, o ribossomo bloqueia fisicamente seu pareamento.

A RNA polimerase continua a se mover ao longo do DNA, consequentemente transcrevendo a região 4 da 5′ UTR. A região 4 é complementar à região 3, e como a região 3 não consegue parear com a região 2, ela pareia com a região 4. O pareamento das regiões 3 e 4 (ver Figura 16.17 H) produz o atenuador, e a transcrição termina logo após a região 4. Os genes estruturais não são transcritos, não são traduzidas as enzimas produtoras de triptofano e não é sintetizado mais triptofano. O **Quadro 16.3** resume os eventos importantes no processo de atenuação.

Um fator-chave para controlar a atenuação é o número de moléculas de tRNA carreadas com triptofano, porque sua disponibilidade é que determina se o ribossomo para nos códons de triptofano. Um segundo fator envolve a sincronização da transcrição e da tradução, que é crítica para a atenuação. A sincronização é alcançada por meio de um sítio de pausa localizado na região 1 da 5′ UTR. Quando esse sítio é transcrito, o RNA se dobra em uma estrutura secundária que inibe a transcrição. Assim, a RNA polimerase para temporariamente no sítio de pausa, dando tempo para o ribossomo se ligar à extremidade 5′ do mRNA. À medida que o ribossomo se aproxima da estrutura secundária no RNA, o ribossomo se divide e então a transcrição continua. Dessa forma, a tradução acompanha de perto a transcrição. É importante destacar que os ribossomos não cruzam os grampos enroscados da 5′ UTR para traduzir os genes estruturais. Os ribossomos que se fixam à extremidade 5′ da região 1 do mRNA encontram um códon de parada na extremidade da região 1. Novos ribossomos que traduzem os genes estruturais se fixam a um diferente local de ligação do ribossomo situado próximo do início do gene *trpE*. **Resolva o Problema 27**

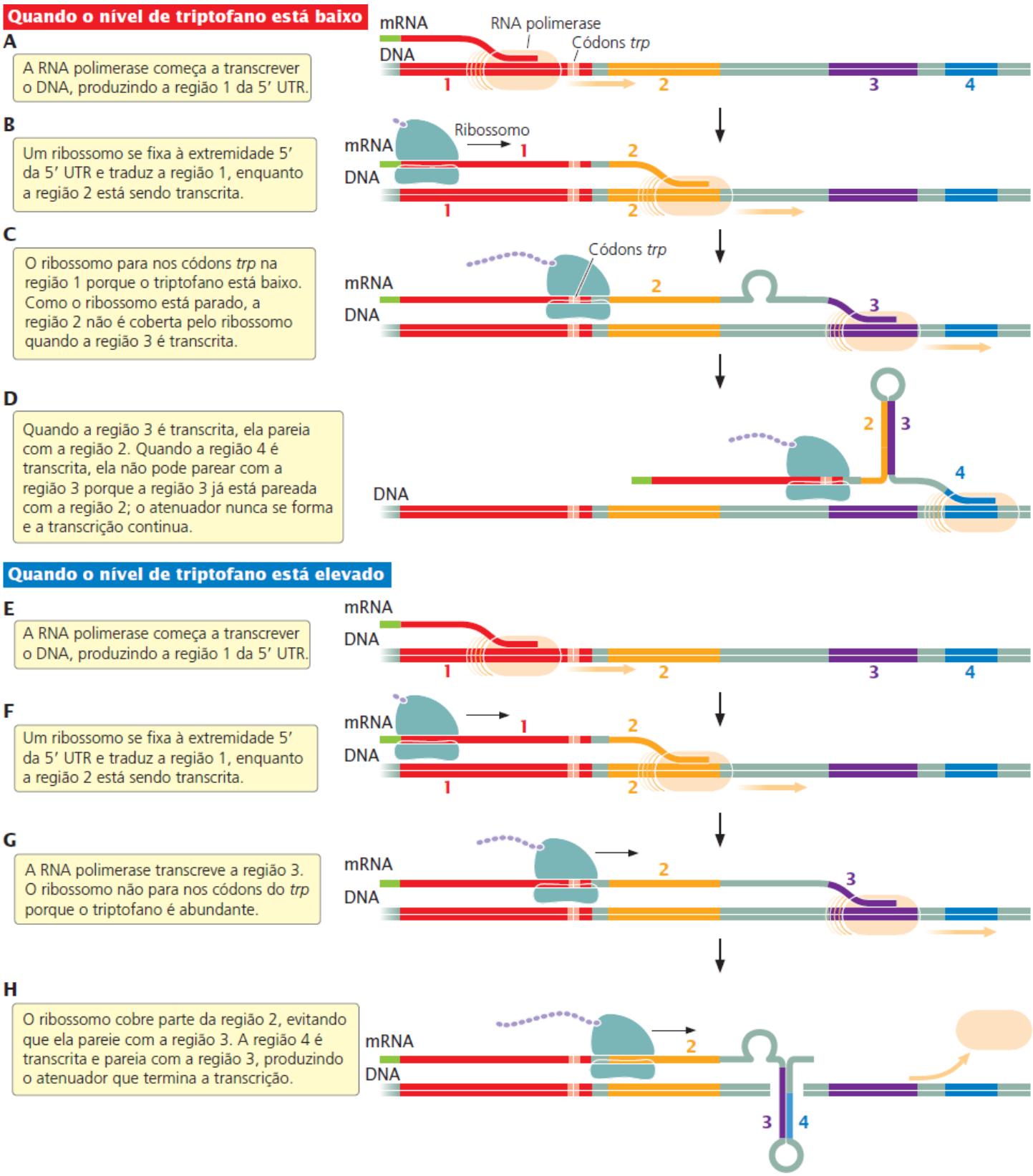


Figura 16.17 **O término prematuro da transcrição (atenuação) ocorre no óperon *trp* dependendo do nível celular de triptofano.**

Quadro 16.3 Eventos no processo de atenuação.

Nível intracelular de triptofano	Ribossomo para nos códons *trp*	Posição do ribossomo quando a região 3 é transcrita	Estrutura secundária da 5′ UTR	Término da transcrição de óperon *trp*
Alta	Não	Cobre a região 2	Grampo 3+4	Sim
Baixa	Sim	Cobre a região 1	Grampo 2+3	Não

Por que ocorre atenuação no óperon *trp*?

Por que as bactérias precisam de atenuação no óperon *trp*? A repressão no sítio do operador não deveria evitar que ocorresse a transcrição quando os níveis de triptofano na célula estão elevados? Por que a célula tem dois tipos de controle? Parte da resposta é que a repressão nunca é completa; parte da transcrição é iniciada mesmo quando o repressor *trp* está ativo; a repressão reduz a transcrição apenas em 70 vezes. A atenuação pode reduzir a transcrição em mais 8 a 10 vezes, então, juntos, os dois processos são capazes de reduzir a transcrição do óperon *trp* em mais de 600 vezes. Ambos os mecanismos fornecem à *E. coli* um controle mais minucioso sobre a síntese do triptofano do que cada um conseguiria sozinho.

Outra resposta para o controle duplo é que a atenuação e a repressão respondem a diferentes sinais: a repressão responde aos níveis celulares de triptofano, enquanto a atenuação responde ao número de tRNAs carreados com triptofano. Podem existir momentos nos quais a capacidade da célula em responder a esses diferentes sinais seja útil. Finalmente, o repressor *trp* afeta vários óperons além do óperon *trp*. Em um estágio inicial da evolução da *E. coli*, o óperon *trp* pode ter sido controlado apenas pela atenuação. O repressor *trp* pode ter evoluído sobretudo para controlar os outros óperons, e afeta apenas incidentalmente o óperon *trp*.

A atenuação é um processo difícil de entender porque é precioso visualizar simultaneamente como dois processos dinâmicos – transcrição e tradução – interagem, e é fácil confundi-los. Lembre-se de que a atenuação se refere ao término precoce da *transcrição*, não da tradução (embora os eventos na tradução levem ao término da transcrição). A atenuação às vezes confunde porque sabemos que a transcrição precisa ocorrer antes da tradução. Estamos confortáveis com a ideia de que a transcrição possa afetar a tradução, mas é mais difícil imaginar que os efeitos da tradução possam influenciar a transcrição, como ocorre na atenuação. A realidade é que a transcrição e a tradução estão intimamente associadas nas células procarióticas, e eventos em um processo podem facilmente afetar o outro.

Conceitos

Na atenuação, a transcrição é iniciada, mas termina de forma prematura. Quando os níveis de triptofano estão baixos, o ribossomo para nos códons do triptofano e a transcrição continua. Quando os níveis de triptofano estão elevados, o ribossomo não para nos códons de triptofano e a 5′ UTR adota uma estrutura secundária que termina a transcrição antes que os genes estruturais possam ser copiados em RNA (atenuação).

Checagem dos conceitos 9

O resultado é a atenuação quando quais regiões da 5′ UTR pareiam?

a. 1 e 3.
b. 2 e 3.
c. 2 e 4.
d. 3 e 4.

16.4 As moléculas de RNA controlam a expressão em alguns genes bacterianos

Todos os reguladores da expressão gênica que analisamos até o momento são proteínas. Também foram descobertos vários exemplos de RNA reguladores.

RNA antissenso

Algumas moléculas de RNA são complementares a sequências específicas em mRNAs e são chamadas de **RNA antissenso**. Elas controlam a expressão gênica ao se ligar a sequências no mRNA e inibir a transcrição. O controle da tradução pelo RNA antissenso é observado na regulação do gene *ompF* da *E. coli* (**Figura 16.18 A**). Esse gene codifica uma proteína da membrana externa que atua como canal para a difusão passiva de pequenas moléculas polares, como água e íons, através da membrana celular. Na maioria das condições, o gene *ompF* é transcrito e traduzido e a proteína OmpF é sintetizada. Entretanto, quando a osmolaridade do meio aumenta, a célula reduz a produção da proteína OmpF para ajudar a manter a osmolaridade celular. Um gene regulador chamado *micF* – abreviação para RNA complementar de interferência de mRNA – é ativado, e o *micF* RNA é produzido (**Figura 16.18 B**). O *micF* RNA, um RNA antissenso, liga-se a uma sequência complementar na 5′ UTR do *ompF* mRNA e inibe a ligação do ribossomo. Essa inibição reduz o volume de tradução (ver Figura 16.18 B), que leva a menos proteínas OmpF na membrana externa e reduz o movimento prejudicial de substâncias através da membrana por causa das

mudanças na osmolaridade. Pesquisa recente encontrou muitas moléculas de RNA antissenso em uma ampla variedade de espécies de bactérias. **Resolva o Problema 30**

Riboswitches

Já observamos que óperons das bactérias têm sequências de DNA (promotores e sítios de operador) onde a ligação de pequenas moléculas induz ou reprime a transcrição. Algumas moléculas de mRNA têm sequências regulatórias chamadas de ***riboswitches*** (ou **ribointerruptores**), onde as moléculas podem se ligar e afetar a expressão gênica ao influenciar a formação de estruturas secundárias no mRNA (**Figura 16.19**). Os *riboswitches* foram descobertos em 2002, e agora parece que são comuns nas bactérias, regulando cerca de 4% de todos os genes bacterianos. Eles também estão são encontrados em archaea, fungos e plantas.

Os *riboswitches* são encontrados na 5′ UTR do mRNA e se dobram em estruturas secundárias de RNA compactas com uma haste de base e vários grampos de ramificação. Em alguns casos, uma pequena molécula regulatória se liga ao *riboswitch* e estabiliza um terminador, que leva ao término prematuro da transcrição. Em outros casos, a ligação de uma molécula regulatória estabiliza uma estrutura secundária que mascara o sítio de ligação do ribossomo, evitando o início da tradução. Quando não está ligado à molécula regulatória, o *riboswitch* assume uma estrutura alternativa que elimina o terminador prematuro ou torna disponível o sítio de ligação ao ribossomo.

Um riboswitch controla a síntese da vitamina B_{12}. Um exemplo de *riboswitch* é observado nos genes bacterianos que codificam as enzimas que participam da síntese da vitamina B_{12}. Os genes para essas enzimas são transcritos em uma molécula de mRNA que tem um *riboswitch*. Quando há a forma ativada da vitamina B_{12} – chamada de coenzima B_{12} –, ela se liga ao *riboswitch* e o mRNA se dobra em uma estrutura secundária que obstrui o sítio de ligação do ribossomo. Consequentemente, não ocorre tradução do mRNA. Não havendo a coenzima B_{12}, o mRNA assume uma estrutura secundária diferente. Essa estrutura secundária não obstrui o sítio de ligação do ribossomo e a tradução inicia, as enzimas são sintetizadas e mais vitamina B_{12} é produzida. Para alguns *riboswitches*, a molécula regulatória atua como um repressor (como já foi descrito) ao inibir a transcrição ou tradução; para outros, a molécula regulatória atua como um indutor, levando à formação de uma estrutura secundária que torna possível a transcrição ou a tradução.

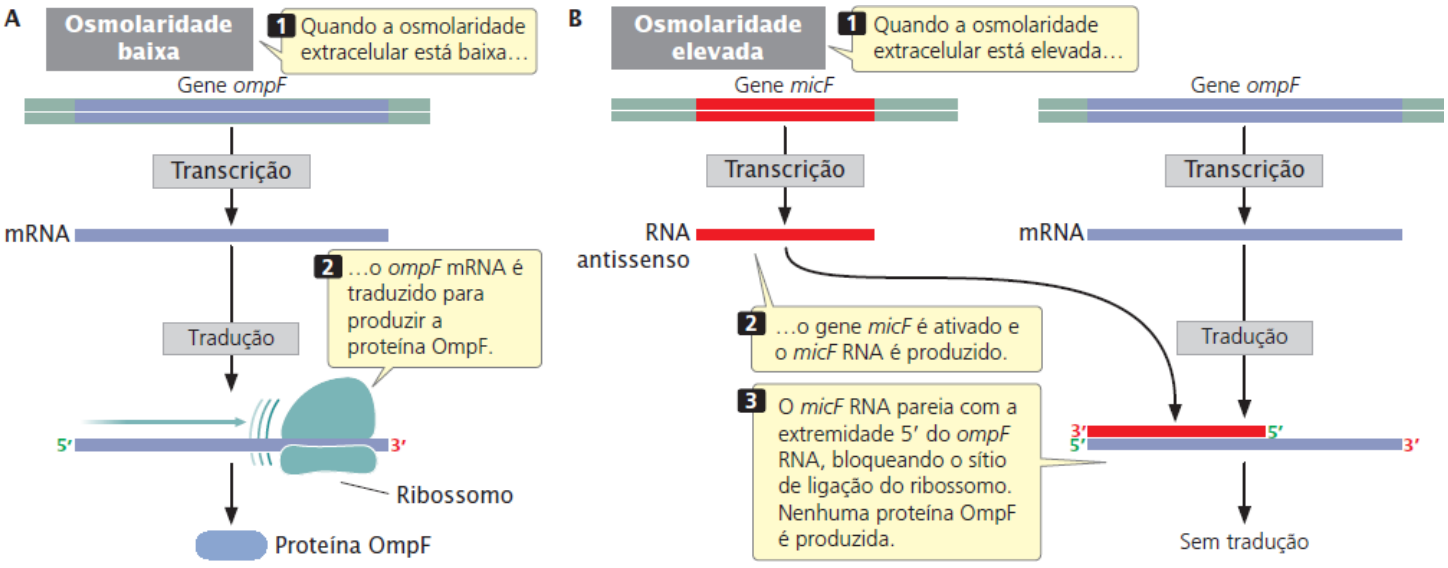


Figura 16.18 O RNA antissenso pode regular a tradução.

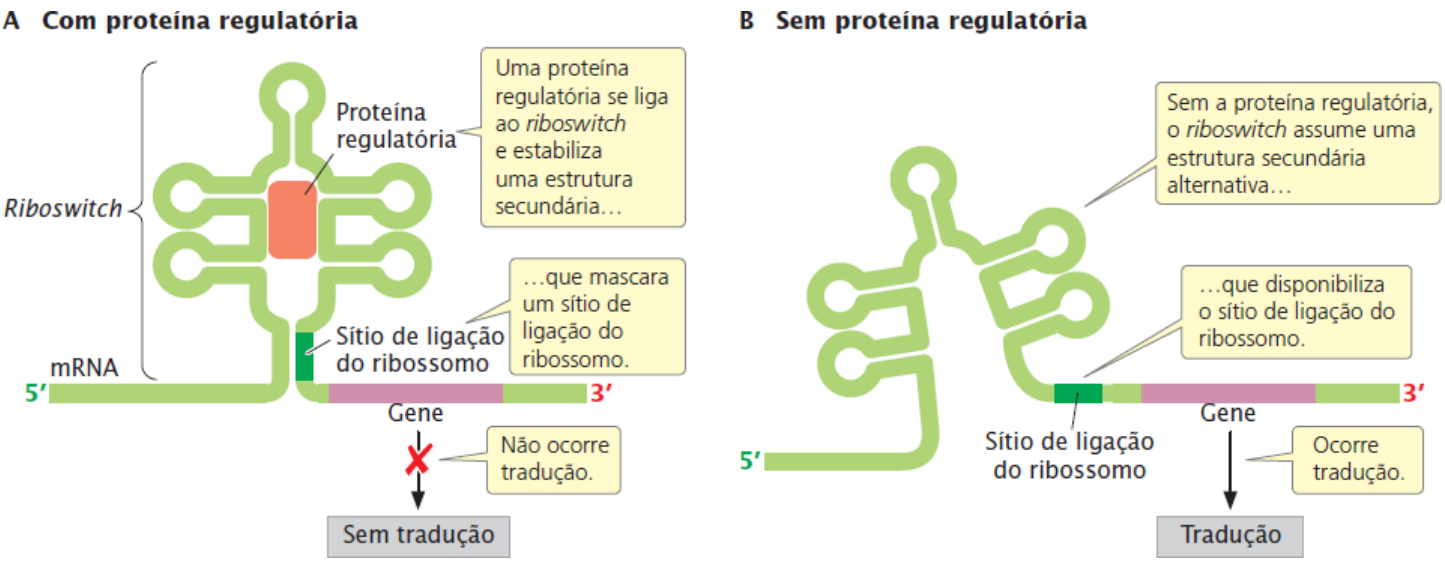


Figura 16.19 **Os *riboswitches* são sequências de RNA no mRNA que afetam a expressão gênica.**

Repressão mediada do RNA por meio de ribozimas

Outro tipo de controle gênico é feito por moléculas de mRNA chamadas ribozimas, que têm atividade catalítica (ver *Capítulo 14*). Chamado de repressão mediada pelo RNA, esse tipo de controle foi demonstrado no gene *glmS* da bactéria *Bacillus subtilis*. A transcrição desse gene produz uma molécula de mRNA que codifica a enzima glutamina-frutose-6-fosfato amidotransferase (**Figura 16.20**), que ajuda a sintetizar um pequeno açúcar chamado glicosamina-6-fosfato (GlcN6P). Existem cerca de 75 nucleotídios dentro da 5′ UTR do mRNA do *glmS* que atuam como uma ribozima. Quando a GlcN6P está ausente, o gene *glmS* é transcrito e traduzido para produzir a enzima, que sintetiza mais GlcN6P. Entretanto, quando há uma quantidade suficiente de GlcN6P, ela se liga à parte ribozima do mRNA, que então induz a autoclivagem do mRNA – a ribozima rompe o esqueleto açúcar fosfato do RNA. Essa ruptura impede a tradução do mRNA.

Conceitos

O RNA antissenso é complementar a outras sequências de RNA ou DNA. Nas células bacterianas, ele pode inibir a tradução ao se ligar a sequências na 5′ UTR do mRNA e evitar a fixação do ribossomo. Os *riboswitches* são sequências nas moléculas de mRNA que se ligam a moléculas regulatórias e induzem mudanças na estrutura secundária do mRNA que afetam a expressão gênica. Na repressão mediada pelo RNA, uma sequência de ribozima no mRNA induz a autoclivagem e a degradação do mRNA quando ligada a uma molécula regulatória.

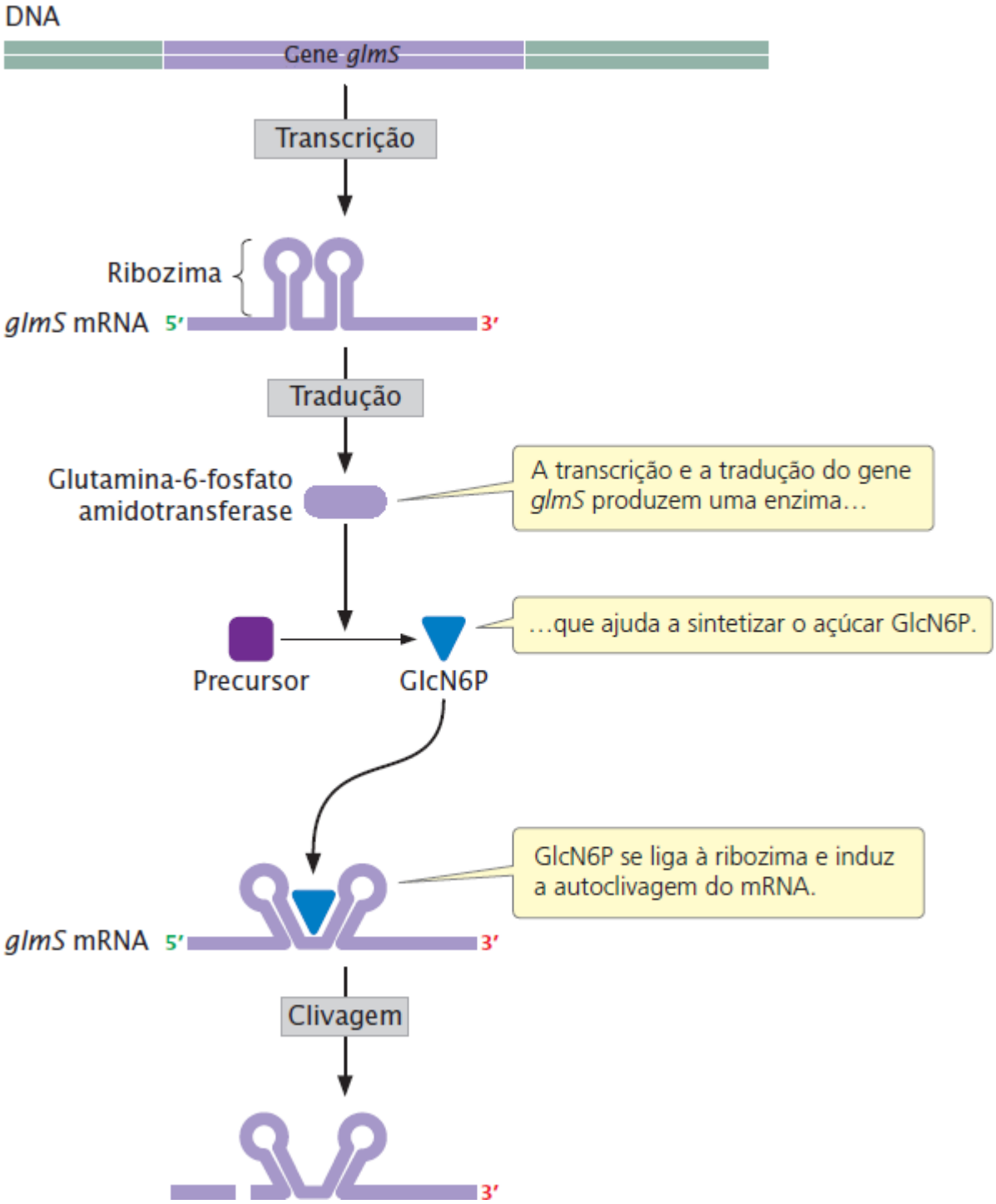


Figura 16.20 **As ribozimas, quando ligadas a pequenas moléculas regulatórias, podem induzir a ruptura e a degradação do mRNA.**

Resumo dos conceitos

- A expressão gênica pode ser controlada em diferentes níveis, incluindo a alteração do DNA ou a estrutura da cromatina, a transcrição, o processamento do mRNA, a estabilidade do RNA, a tradução e a modificação pós-traducionais. Boa parte da regulação gênica é feita por meio da ação de proteínas regulatórias que se ligam a sequências específicas no DNA
- Os genes nas células bacterianas estão agrupados em óperons – grupos de genes estruturais relacionados por sua função e sequências que controlam sua transcrição. Os genes estruturais em um óperon são transcritos juntos em uma única molécula de mRNA
- No controle negativo, uma proteína repressora se liga ao DNA e inibe a transcrição. No controle positivo, uma proteína ativadora se liga ao DNA e estimula a transcrição. Nos óperons induzíveis, a transcrição está desligada e deve ser ligada; nos óperons reprimíveis, a transcrição está ligada e deve ser desligada
- O óperon *lac* da *E. coli* é um óperon induzível negativo. Sem a lactose, um repressor se liga ao operador e evita a transcrição de genes que codificam β-galactosidase, permease e transacetilase. Quando há lactose, parte dela é convertida em alolactose,

que se liga ao repressor e o inativa, possibilitando que os genes estruturais sejam transcritos

- O controle positivo no *lac* e em outros óperons é feito por meio da repressão catabólica. Quando complexada com o AMP cíclico, a proteína ativadora catabólica se liga ao sítio ou próximo ao promotor e estimula a transcrição dos genes estruturais. Os níveis de cAMP estão inversamente relacionados com a glicose; então, níveis baixos de glicose estimulam a transcrição e níveis elevados inibem a transcrição
- O óperon *trp* da *E. coli* é um óperon reprimível negativo que controla a biossíntese do triptofano
- A atenuação provoca o término prematuro da transcrição. Ela ocorre graças ao acoplamento da transcrição e tradução e depende da estrutura secundária da sequência 5′ UTR
- Os RNAs antissenso são complementares a sequências no mRNA e inibem a tradução ao se ligar a essas sequências, evitando a fixação ou o progresso do ribossomo
- Quando ligados a uma molécula regulatória, os *riboswitches* nas moléculas de mRNA induzem mudanças na estrutura secundária do mRNA, que afetam a expressão gênica. Alguns mRNAs têm sequências de ribozima que induzem a autoclivagem e a degradação quando ligados a uma molécula regulatória.

Termos importantes

Antiterminadora
Atenuação
Atenuador
Controle negativo
Controle positivo
Correpressor
Diploide parcial
Domínio
Elemento regulador
Gene constitutivo
Gene estrutural
Gene regulador
Gene regulatório
Indução coordenada
Indutor
Monofosfato de adenosina-3′,5′-cíclico (cAMP)
Operador
Óperon
Óperon induzível
Óperon reprimível
Proteína alostérica
Proteína catabólica ativadora (CAP)
Proteína reguladora
Regulação gênica
Repressão catabólica
Riboswitch
RNA antissenso

Respostas da Checagem dos conceitos

1. Um gene constitutivo não é regulado e é expresso de forma contínua.
2. Porque ela é a primeira etapa no processo de transferência de informação do DNA para proteína. Para eficiência celular, a expressão gênica é, com frequência, regulada no início no processo de produção de proteína.
3. b.
4. d.
5. d.
6. c.
7. b.
8. b.

9. d.

Problemas desenvolvidos

Problema 1

O óperon *fox*, que tem as sequências *A, B, C* e *D* (que podem representar os genes estruturais ou sequências regulatórias), codifica as enzimas 1 e 2. As mutações nas sequências *A*, *B*, *C* e *D* têm os seguintes efeitos, em que um sinal de mais (+) indica que a enzima é sintetizada e um sinal de menos (−) indica que a enzima não é sintetizada.

	Sem fox		Com fox	
Mutação na sequência	**Enzima 1**	**Enzima 2**	**Enzima 1**	**Enzima 2**
Sem mutação	−	−	+	+
A	−	−	−	+
B	−	−	−	−
C	−	−	+	−
D	+	+	+	+

a. O óperon *fox* é induzível ou reprimível?
b. Indique qual sequência (*A*, *B*, *C* ou *D*) é parte dos seguintes componentes do óperon. Cada sequência deve ser usada apenas uma vez.

Gene regulador _____
Promotor _____
Gene estrutural para enzima 1 _____
Gene estrutural para enzima 2 _____

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

a. Se o óperon fox é induzível ou reprimível.
b. Qual sequência representa cada parte do óperon.

Quais informações são fornecidas para solucionar o problema?

Para cada mutação, se a enzima 1 ou a enzima 2 é produzida, havendo ou não Fox.

Para a solução deste problema, revise:

Seção 16.2.

Etapas para a solução

a. Quando não existem mutações, as enzimas 1 e 2 são produzidas quando há Fox, mas não quando ele está ausente, indicando que o óperon é induzível e que o Fox é o indutor.
b. A mutação em *A* possibilita a produção da enzima 2 quando há Fox, mas a enzima 1 não é produzida quando há, nem quando não há Fox e, então, *A* deve ter uma mutação no gene estrutural para enzima 1. Com a mutação em *B*, nenhuma enzima é produzida sob nenhuma condição, e então é provável que essa mutação ocorra no promotor e impeça a RNA polimerase de se ligar. A mutação em *C* afeta apenas a enzima 2, que não é produzida na presença ou ausência de Fox; a enzima 1 é produzida normalmente (apenas quando há Fox), e é provável que a mutação em *C* ocorra no gene estrutural para enzima 2. A mutação em *D* é constitutiva, possibilitando a produção das enzimas 1 e 2 independente de haver Fox. É provável que essa mutação ocorra no gene regulador, produzindo um repressor defeituoso que é incapaz

de se ligar ao operador sob nenhuma circunstância.

Dica: Revise a Figura 16.3 sobre o resumo da estrutura do óperon.

Gene regulador *D* Gene estrutural para enzima 1 *A*
Promotor *B* Gene estrutural para enzima 2 *C*

Problema 2

Uma mutação ocorre na 5′ UTR do óperon *trp* que reduz a capacidade de a região 2 parear com a região 3. Qual será o efeito dessa mutação quando o nível de triptofano está elevado? Quando o nível de triptofano está baixo?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
O efeito da mutação quando triptofano está elevado e quando ele está baixo.

Quais informações são fornecidas para solucionar o problema?
- A mutação ocorre na 5′ UTR do óperon *trp*
- A mutação reduz a capacidade da região 2 em parear com a região 3.

Para a solução deste problema, revise:
Seção 16.3.

Etapas para a solução

Quando o nível de triptofano está elevado, as regiões 2 e 3 não pareiam normalmente e, portanto, a mutação não terá efeito. Quando o nível de triptofano está baixo, entretanto, o ribossomo normalmente para nos códons de triptofano na região 1 e não cobre a região 2, e as regiões 2 e 3 ficam livres para parear, o que evita que as regiões 3 e 4 pareiem e formem um terminador, que encerraria a transcrição. Se as regiões 2 e 3 não conseguirem parear, as regiões 3 e 4 vão parear mesmo quando o triptofano estiver baixo, e a atenuação sempre acontecerá. Portanto, não será mais sintetizado triptofano mesmo na sua ausência.

Dica: Revise a Figura 16.17 para ler um resumo sobre a atenuação.

Questões de compreensão

Seção 16.1

1. Por que a expressão gênica é importante para as células bacterianas?
2. Cite seis níveis diferentes nos quais a expressão gênica pode ser controlada.

Seção 16.2

3. Faça um desenho ilustrando a estrutura geral de um óperon e identifique suas partes.
4. Qual a diferença entre o controle positivo e negativo? Qual a diferença entre os óperons induzíveis e reprimíveis?
5. Descreva resumidamente o óperon *lac* e como ele controla o metabolismo da lactose.
6. O que é repressão catabólica? De que modo ela age para que uma célula bacteriana use glicose em detrimento de outros açúcares?

Seção 16.3

7. O que é atenuação? Quais são os mecanismos pelos quais o atenuador se forma quando os níveis de triptofano estão elevados e o antiterminador se forma quando os níveis de triptofano estão baixos?

Seção 16.4

8. O que é RNA antissenso? Como ele controla a expressão gênica?
9. O que são *riboswitches*? Como eles controlam a expressão gênica? Qual é a diferença dos *riboswitches* para a repressão mediada pelo RNA?

Questões e problemas aplicados

Seção 16.1

10. Examine a **Figura 16.2 B**. Por que você acha que o motivo da proteína ligadora de DNA apresentado é chamado de proteína de dedo-de-zinco?

Seção 16.2

*11. Para cada um dos tipos de controle de transcrição a seguir, indique se a proteína produzida pelo gene regulador será sintetizada inicialmente como um repressor ativo, repressor inativo, ativador ativo ou ativador inativo.
    a. Controle negativo em um óperon reprimível.
    b. Controle positivo em um óperon reprimível.
    c. Controle negativo em um óperon induzível.
    d. Controle positivo em um óperon induzível.
12. Uma mutação no sítio do operador evita a ligação da proteína reguladora. Que efeito essa mutação terá nos seguintes tipos de óperons?
    a. A proteína reguladora é um repressor em um óperon reprimível.
    b. A proteína reguladora é um repressor em um óperon induzível.
13. O óperon *blob* produz enzimas que convertem o composto A no composto B. O óperon é controlado por um gene regulatório *S*. Em condições normais, as enzimas são sintetizadas apenas na ausência do composto B. Se o gene *S* sofrer uma mutação, as enzimas serão sintetizadas havendo ou não o composto B. O gene *S* produz um repressor ou um ativador? Esse óperon é induzível ou reprimível?
*14. Uma mutação impede que a proteína catabólica ativadora (CAP) se ligue ao promotor no óperon *lac*. Qual será o efeito dessa mutação na transcrição do óperon?
15. Transformação é um processo no qual as bactérias captam DNA novo liberado por células mortas e o integram aos próprios genomas (ver *Capítulo 9*). Na infecção por *Streptococcus pneumoniae* (responsável por muitos casos de pneumonia, otite média e meningite), a capacidade de fazer a transformação requer de 105 a 124 genes, coletivamente chamados de régulon *com*. O régulon *com* é ativado em resposta a uma proteína chamada peptídio estimulante de competência (CSP), que é produzida pelas bactérias e exportada para o meio circundante. Quando se acumula CSP suficiente, ele se fixa a um receptor na membrana celular da bactéria, que então ativa uma proteína reguladora que estimula a transcrição de genes no régulon *com* e desencadeia reações que, por fim, resultam em transformação. O régulon *com* na bactéria *Streptococcus pneumoniae* parece ser controlado por qual tipo de regulação gênica? Explique sua resposta.
    a. Induzível negativo.
    b. Reprimível negativo.
    c. Induzível positivo.
    d. Reprimível positivo.
16. Em quais das seguintes condições o óperon *lac* produziria a maior quantidade de *β*-galactosidase? E a menor? Explique sua resposta.

	Com lactose	**Com glicose**
Condição 1	Sim	Não
Condição 2	Não	Sim
Condição 3	Sim	Sim
Condição 4	Não	Não

17. Uma cepa mutante de *E. coli* produz *β*-galactosidase quando há e quando não há lactose. Onde ocorre a mutação no óperon nessa cepa?
18. Examine a Figura 16.8. Qual seria o efeito de um fármaco que modificasse a estrutura de alolactose de modo que ela não conseguisse se ligar à proteína reguladora?
*19. Para cepas de *E. coli* com os genótipos *lac* apresentados a seguir, use um sinal de mais (+) para indicar a síntese de *β*-galactosidase e permease e um sinal de menos (−) para indicar ausência de síntese das proteínas.

	Sem lactose		**Com lactose**	
	***β*-**	**Permease**	***β*-**	**Permease**

Genótipo da cepa	galactosidase	galactosidase
lacI$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^+$		
lacI$^-$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^+$		
lacI$^+$ *lacP*$^+$ *lacO*^{c} *lacZ*$^+$ *lacY*$^+$		
lacI$^-$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^-$		
lacI$^-$ *lacP*$^-$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^+$		
lacI$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^-$ *lacY*$^+$/ *lacI*$^-$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^-$		
lacI$^-$ *lacP*$^+$ *lacO*^{c} *lacZ*$^+$ *lacY*$^+$/ *lacI*$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^-$ *lacY*$^-$		
lacI$^-$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^-$/ *lacI*$^+$ *lacP*$^-$ *lacO*$^+$ *lacZ*$^-$ *lacY*$^+$		
lacI$^+$ *lacP*$^-$ *lacO*^{c} *lacZ*$^-$ *lacY*$^+$/ *lacI*$^-$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^-$		
lacI$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^+$/ *lacI*$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^+$		
lacI^s *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^-$/ *lacI*$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^-$ *lacY*$^+$		
lacI^s *lacP*$^-$ *lacO*$^+$ *lacZ*$^-$ *lacY*$^+$/ *lacI*$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$ *lacY*$^+$		

20. Indique todos os possíveis genótipos de um óperon *lac* que produz *β*-galactosidase e permease sob as condições a seguir. Não indique os genótipos diploides parciais.

	Sem lactose		Com lactose	
	***β*-galactosidase**	**Permease**	***β*-galactosidase**	**Permease**
a.	−	−	+	+
b.	−	−	−	+
c.	−	−	+	−
d.	+	+	+	+
e.	−	−	−	−
f.	+	−	+	−
g.	−	+	−	+

*21. Explique por que as mutações no gene *lacI* são trans em seus efeitos, mas as mutações no gene *lacO* são cis em seus efeitos.

22. Qual fita de DNA (superior ou inferior) na Figura 16.9 é a fita molde? Explique sua resposta.

23. O óperon *mmm*, que tem as sequências *A*, *B*, *C* e *D* (que podem representar os genes estruturais ou sequências regulatórias), codifica as enzimas 1 e 2. As mutações nas sequências *A*, *B*, *C* e *D* têm os seguintes efeitos, em que um sinal de mais (+) indica que a enzima é sintetizada e um sinal de menos (−) indica que a enzima não é sintetizada.

Mutação na sequência	Sem *mmm*		Com *mmm*	
	Enzima 1	Enzima 2	Enzima 1	Enzima 2
Sem mutação	+	+	−	−
A	−	+	−	−
B	+	+	+	+
C	+	−	−	−
D	−	−	−	−

a. O óperon *mmm* é induzível ou reprimível?
b. Indique qual sequência (*A*, *B*, *C* ou *D*) é parte dos seguintes componentes do óperon:

Gene regulador _____
Promotor _____
Gene estrutural para enzima 1 _____
Gene estrutural para enzima 2 _____

ANÁLISE DE DADOS

24. Ellis Engelsberg e seus colaboradores examinaram a regulação dos genes que participam do metabolismo da arabinose, um açúcar (E. Engelsberg *et al*. 1965. *Journal of Bacteriology* 90:946-957). Quatro genes estruturais codificam as enzimas que ajudam a metabolizar a arabinose (genes *A*, *B*, *D* e *E*). Um gene *C* adicional é ligado aos genes *A*, *B* e *D*. Esses genes estão na ordem *D*-*A*-*B*-*C*. O gene *E* está distante dos outros. Engelsberg e seus colegas isolaram mutações no gene *C* que afetavam a expressão dos genes estruturais *A*, *B*, *D* e *E*. Em um conjunto de experimentos, eles criaram vários genótipos nos *loci A* e *C* e determinaram se a arabinose isomerase (a enzima codificada pelo gene *A*) era produzida havendo ou não arabinose (o substrato da arabinose isomerase). O quadro a seguir mostra os resultados desse experimento, em que um sinal de mais (+) indica que a enzima é sintetizada e um sinal de menos (−) indica que a enzima não é sintetizada.

Genótipo	Sem arabinose	Com arabinose
1. $C^+ A^+$	−	+
2. $C^- A^+$	−	−
3. $C^- A^+/C^+A^-$	−	+
4. $C^c A^- /C^- A^+$	+	+

a. Com base nos resultados desses experimentos, o gene C é um gene operador ou regulador? Explique sua resposta.
b. Esses experimentos sugerem que o óperon arabinose sofria controle negativo ou positivo? Explique sua resposta.
c. Qual tipo de mutação é C^c?

ANÁLISE DE DADOS

25. Na *E. coli*, três genes estruturais (*A*, *D* e *E*) codificam as enzimas A, D e E, respectivamente. O gene *O* é um operador. Os genes estão na ordem *O*-*A*-*D*-*E* no cromossomo. Essas enzimas catalisam a biossíntese de valina. Foram isoladas mutações nos genes *A*, *D*, *E* e *O* para estudar a produção das enzimas A, D e E quando os níveis celulares de valina estiverem baixos (T. Ramakrishnan and E. A. Adelberg. 1965. *Journal of Bacteriology* 89:654-660). O quadro a seguir mostra os níveis das enzimas produzidas pela *E. coli* diploide parcial com várias combinações de mutações.

Genótipo	Quantidade de enzima produzida		
	E	D	A
1. $E^+ D^+ A^+ O^+/ E^+ D^+ A^+ O^+$	2,40	2,00	3,50
2. $E^+ D^+ A^+ O^-/$	35,80	38,60	46,80

$E^+ D^+ A^+ O^+$			
3. $E^+ D^- A^+ O^-$/ $E^+ D^+ A^- O^+$	1,80	1,00	47,00
4. $E^+ D^+ A^- O^-$/ $E^+ D^- A^+ O^+$	35,30	38,00	1,70
5. $E^- D^+ A^+ O^-$/ $E^+ D^- A^+ O^+$	2,38	38,00	46,70

a. A proteína reguladora que se liga ao operador desse óperon é repressora (controle negativo) ou ativadora (controle positivo)? Explique sua resposta.
b. Os genes *A*, *D* e *E* estão todos sob o controle do operador *O*? Explique sua resposta.
c. Proponha uma explicação para o nível baixo da enzima E produzida no genótipo 3.

Seção 16.3

26. Em que nível da regulação gênica apresentada na **Figura 16.1** ocorre a atenuação?

*27. Algumas mutações encontradas na região 5′ UTR do óperon *trp* da *E. coli* estão listadas de *a* a *g*. Qual é o efeito mais provável de cada uma dessas mutações na transcrição dos genes estruturais *trp*?
a. Uma mutação que impede a ligação do ribossomo à extremidade 5′ UTR do mRNA.
b. Uma mutação que muda os códons do triptofano na região 1 da 5′ UTR do mRNA em códons para alanina.
c. Uma mutação que cria um códon de parada na região 1 da 5′ UTR do mRNA.
d. Deleções na região 2 da 5′ UTR do mRNA.
e. Deleções na região 3 da 5′ UTR do mRNA.
f. Deleções na região 4 da 5′ UTR do mRNA.
g. Deleção de um anel de nucleotídios de adenina que segue a região na 5′ UTR.

28. Algumas mutações na região 5′ UTR aumentam o término pelo atenuador. Onde essas mutações poderiam ocorrer e como elas poderiam afetar o atenuador?

29. Algumas das mutações mencionadas no Problema 28 têm uma propriedade interessante. Elas impedem a formação do antiterminador que normalmente ocorre quando o nível de triptofano está baixo. Em uma das mutações, o códon de início AUG para o peptídio 5′ UTR foi eliminado. Como essa mutação poderia impedir que ocorra o antitérmino?

Seção 16.4

*30. Foram descobertos vários exemplos de RNA antissenso que regulam a tradução nas células bacterianas. Os geneticistas moleculares também usaram o RNA antissenso para controlar artificialmente a transcrição nos genes bacterianos e eucarióticos. Se você quiser inibir a transcrição de um gene bacteriano com o RNA antissenso, quais sequências esse RNA antissenso teria?

Questão desafiadora

Seção 16.3

31. Você esperaria ver atenuação no óperon *lac* e em outros óperons que controlam o metabolismo dos açúcares? Por que sim ou por que não?

17 Controle da Expressão Gênica nos Eucariotos

As diferenças genéticas que nos tornam humanos

Há mais de 140 anos, Charles Darwin propôs que os humanos compartilhavam um ancestral recente comum com grandes primatas africanos. Atualmente, existem evidências significativas de que nosso parente vivo mais próximo seja o chimpanzé. Fósseis indicam que os seres humanos e os chimpanzés se separaram geneticamente há apenas 5 a 7 milhões de anos – um simples piscar de olhos na escala evolutiva. Os seres humanos e chimpanzés apresentam diferenças em muitos traços anatômicos, fisiológicos, comportamentais e cognitivos. Por exemplo, há numerosas diferenças na estrutura da coluna vertebral, da pelve, do crânio, da mandíbula, dos dentes, das mãos e dos pés dos humanos e dos chimpanzés. O tamanho do cérebro humano é mais de duas vezes o cérebro do chimpanzé, e os humanos exibem uma linguagem complexa e características culturais não observadas nos chimpanzés. Na verdade, o grau de diferença fenotípica entre chimpanzés e humanos é tão grande que os cientistas os colocam em famílias de primatas totalmente diferentes (os humanos na família Hominidae e os chimpanzés na família Pongidae).

Apesar do grande abismo fenotípico entre humanos e chimpanzés, o sequenciamento dos seus genomas revela que o DNA é muito semelhante. Apenas 1% dos pares de bases individuais apresenta diferença entre as duas espécies, junto com uma diferença de 3% nas inserções e deleções. Assim, 96% do DNA dos humanos e chimpanzés são idênticos. No entanto, é nítido que os humanos não são chimpanzés. Como, então, humanos e chimpanzés vieram a ser tão diferentes? Onde estão os genes que nos tornam humanos?

Uma possível resposta a esse paradoxo foi proposta pelos geneticistas Mary-Claire King e A. C. Wilson, em 1975. Usando as limitadas técnicas disponíveis naquela época (comparação dos aminoácidos em proteínas e estudos de hibridização do DNA), King e Wilson concluíram que os humanos e chimpanzés apresentam uma diferença de apenas 1% em suas sequências de DNA. Para explicar como essas mudanças genéticas tão pequenas são responsáveis pelas grandes diferenças físicas e comportamentais entre humanos e chimpanzés, King e Wilson sugeriram que as variações que nos tornam humanos estejam concentradas nas sequências regulatórias – as partes do genoma que controlam a expressão de outros genes. Dessa maneira, pequenas mudanças genéticas poderiam influenciar a expressão de numerosos genes e influenciariam os fenótipos de muitos traços simultaneamente. Infelizmente, não existiam técnicas disponíveis naquele momento para examinar as sequências regulatórias e testar essa hipótese.

Avançando para 2009. Usando tecnologia de ponta de pesquisa genômica e bioinformática, Katja Nowick e seus colegas na University of Illinois e na Norwegian University of Science and Technology identificaram um grupo de 90 fatores de transcrição cuja expressão era muito diferente em humanos e chimpanzés. Como foi discutido no Capítulo 13, os fatores de transcrição são proteínas que se ligam ao DNA e facilitam ou reprimem a síntese de RNA, a primeira etapa no processo de transferência de informação do genótipo para o fenótipo. Cada fator de transcrição pode afetar a expressão de múltiplos genes, então uma pequena mudança genética que afete a expressão de um único fator de transcrição pode influenciar muitos genes adicionais. As diferenças que Nowick e seus colegas encontraram na expressão dos fatores de transcrição eram evidentes sobretudo no tecido cerebral, onde seriam responsáveis pelas grandes diferenças nas funções neurológica e cognitiva de humanos e chimpanzés.

Muitos dos fatores de transcrição que Nowick e seus colegas identificaram eram proteínas de dedo-de-zinco com domínio associado ao módulo (*box*) Krüppel (KRAB-ZFPs), fatores de transcrição que se ligam a sequências específicas de DNA e provocam mudanças na estrutura da cromatina. Como discutiremos neste capítulo, alterações na estrutura da cromatina estão envolvidas na regulação da expressão gênica. Outros estudos demonstraram que KRAB-ZFPs evoluíram rapidamente nos humanos, provavelmente porque foram favorecidos pela seleção natural. Os fatores de transcrição identificados no estudo de Nowick se agruparam em duas redes regulatórias distintas, mas interconectadas, que controlam o metabolismo de energia, a transcrição, o transporte vesicular e a modificação de proteínas.

Esses resultados apoiam a ideia proposta inicialmente por King e Wilson: as mudanças em um número relativamente pequeno de sequências afetam a expressão de numerosos genes nos humanos e nos chimpanzés e ajudam a produzir as grandes diferenças observadas na anatomia, no tamanho do cérebro, na cognição e no comportamento.

Este capítulo discute a regulação gênica nas células eucarióticas, as várias mudanças que ajudam os humanos a serem únicos. A regulação gênica ocorre tipicamente em múltiplos níveis. Inicialmente analisaremos como a expressão gênica é influenciada por mudanças na estrutura da cromatina, que pode ser alterada por vários mecanismos diferentes. A seguir, analisaremos o início da transcrição nos eucariotos, que é controlado por uma interação complexa de vários tipos de proteínas e os elementos regulatórios de DNA aos quais elas se ligam. Examinaremos as várias maneiras pelas quais a expressão gênica é controlada por meio do processamento, da degradação e da tradução do mRNA. Por fim, revisaremos algumas das semelhanças e diferenças na regulação gênica nas bactérias e nos eucariotos.

17.1 Células eucarióticas e bactérias têm muitas características de regulação gênica em comum, mas existem diferenças importantes

Como foi discutido no Capítulo 16, muitas características de regulação gênica são comuns nas células bacterianas e eucarióticas. Por exemplo, em ambos os tipos de células, as proteínas ligadoras de DNA influenciam a capacidade da RNA polimerase em iniciar a transcrição. Entretanto, também existem algumas diferenças. Primeiro, muitos genes procarióticos estão organizados em óperons e são transcritos em uma única molécula de RNA. Embora tenham sido encontrados alguns *clusters* de genes semelhantes ao óperon em vermes e até em alguns cordados primitivos, a maioria dos genes eucarióticos tem seus próprios promotores e é transcrita separadamente. Segundo, a estrutura da cromatina afeta a expressão gênica nas células eucarióticas; o DNA precisa se desenroscar a partir das proteínas histona antes de iniciar a transcrição. Terceiro, a membrana nuclear nas células eucarióticas separa a transcrição e a tradução no tempo e no espaço. Portanto, a regulação da expressão gênica nas células eucarióticas é caracterizada por maior diversidade de mecanismos que atuam em pontos diferentes na transferência de informações do DNA para as proteínas.

A regulação gênica eucariótica é menos compreendida do que a bacteriana, em parte por causa dos genomas maiores nos eucariotos, de sua maior complexidade das sequências e da dificuldade em isolar e manipular mutações que possam ser usadas no estudo. Contudo, houve grandes avanços em nossa compreensão da regulação dos genes eucarióticos nos últimos anos.

17.2 Alterações na estrutura da cromatina influenciam a expressão dos genes

Um tipo de controle gênico nas células eucarióticas é feito pela modificação da estrutura da cromatina. No núcleo, as proteínas histona se associam e formam octâmeros, ao redor dos quais o DNA helicoidal se enrosca firmemente para criar a cromatina (ver Figura 11.3). No senso geral, essa estrutura de cromatina reprime a expressão gênica. Para um gene ser transcrito, os fatores de transcrição, outras proteínas regulatórias e RNA polimerase precisam se ligar ao DNA. Como esses eventos podem ocorrer com o DNA firmemente enroscado ao redor das proteínas histona? A resposta é que, antes da transcrição, a estrutura da cromatina muda e o DNA se torna mais acessível para o maquinário de transcrição.

Hipersensibilidade da DNase I

São observados vários tipos de mudanças na estrutura da cromatina quando os genes se tornam *ativos* para a transcrição. À medida que isso acontece, as regiões ao redor dos genes se tornam muito sensíveis à ação da DNase I (ver *Capítulo 11*). Essas regiões, chamadas de **sítios hipersensíveis da DNase I**, frequentemente se desenvolvem cerca de 1.000 nucleotídios *upstream* do sítio de início da transcrição, sugerindo que a cromatina nessas regiões adota uma configuração mais aberta durante a transcrição. Esse relaxamento da estrutura da cromatina possibilita que as proteínas regulatórias acessem os sítios de ligação ao DNA. De fato, muitos sítios hipersensíveis da DNase I correspondem aos sítios de ligação conhecidos para as proteínas regulatórias. Pelo menos três processos diferentes afetam a regulação gênica ao alterar a estrutura da cromatina: (1) remodelagem da cromatina; (2) modificação das proteínas histona; e (3) metilação do DNA. Cada um desses mecanismos será discutido nas seções a seguir.

Conceitos

A sensibilidade à digestão pela DNase I indica que o DNA transcrito adota uma configuração aberta antes da transcrição.

Remodelagem da cromatina

Alguns fatores de transcrição e outras proteínas regulatórias alteram a estrutura da cromatina sem alterar diretamente a estrutura química das histonas. Essas proteínas são chamadas de **complexos de remodelagem da cromatina**. Elas se ligam diretamente a sítios específicos no DNA e reposicionam os nucleossomos, possibilitando que os fatores de transcrição e a RNA polimerase se liguem aos promotores e iniciem a transcrição (**Figura 17.1**).

Um dos exemplos mais bem estudados de complexo de remodelagem da cromatina é o SWI-SNF, encontrado nas leveduras, nos humanos, na *Drosophila* e em outros organismos. Esse complexo utiliza energia derivada da hidrólise do ATP para reposicionar os nucleossomos, expondo os promotores no DNA à ação de outras proteínas regulatórias e da RNA polimerase.

As evidências sugerem pelo menos dois mecanismos pelos quais os complexos de remodelagem reposicionam os nucleossomos. Primeiro, alguns complexos de remodelagem fazem com que o nucleossomo deslize pelo DNA, criando condições para que o DNA enroscado no nucleossomo ocupe uma posição entre os nucleossomos, onde é mais acessível para as proteínas que afetam a expressão gênica (ver Figura 17.1). Segundo, alguns complexos provocam mudanças conformacionais no DNA, nos nucleossomos ou em ambos, de modo que o DNA ligado ao nucleossomo adote uma configuração mais exposta.

Os complexos de remodelagem da cromatina são direcionados para sequências de DNA específicas por meio de ativadores ou repressores de transcrição que se fixam a um complexo de remodelagem e, então, se ligam aos promotores de genes específicos. Também há indícios de que os complexos de remodelagem da cromatina trabalham em conjunto com enzimas que modificam as histonas, como as enzimas acetiltransferase (ver a próxima seção), para mudar a estrutura da cromatina e expor o DNA para transcrição.

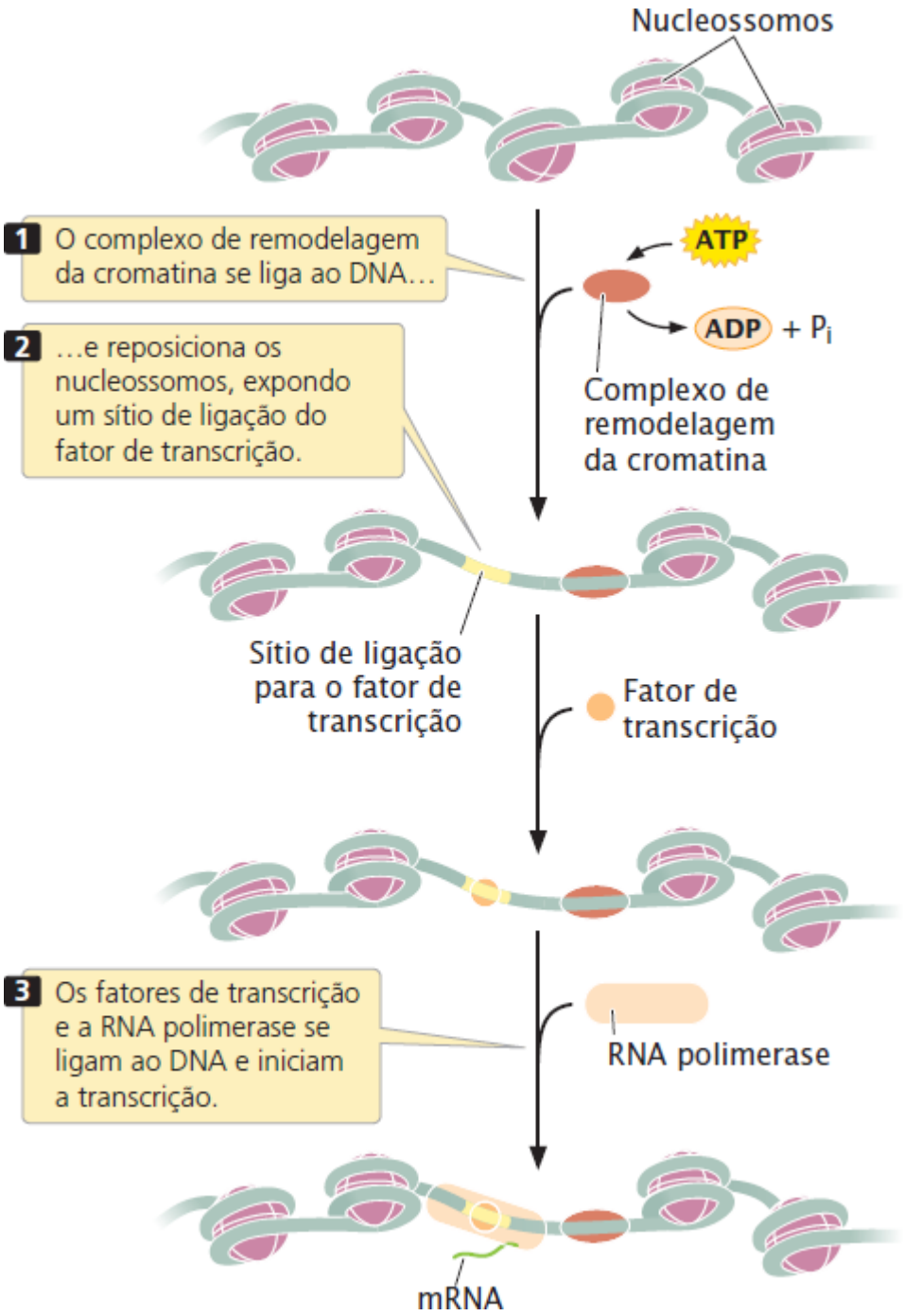


Figura 17.1 Os complexos de remodelagem da cromatina reposicionam os nucleossomos, permitindo que os fatores de transcrição e a RNA polimerase se liguem aos promotores e iniciem a transcrição.

Modificação das histonas

As histonas no octâmero central do nucleossomo têm dois domínios: (1) um domínio globular que se associa com outras histonas e o DNA e (2) um domínio de cauda com carga elétrica positiva que interage com os grupos fosfato com carga elétrica negativa no esqueleto do DNA. As caudas das proteínas histonas são modificadas com a adição ou a remoção de grupos fosfato, metila ou acetila. Outra modificação das histonas é a ubiquitinação, na qual pequenas moléculas chamadas ubiquitina são adicionadas ou removidas das histonas. Todas essas modificações são, às vezes, chamadas de **código da histona**, porque codificam informações que afetam a maneira como os genes são expressos. O código da histona afeta a expressão gênica ao modificar a estrutura da cromatina diretamente, ou, em alguns casos, ao servir como sítios de reconhecimento para proteínas que se ligam ao DNA e então regulam a transcrição.

Metilação das histonas. Um tipo de modificação da histona é a adição de grupos metila às caudas das proteínas histona. Essas modificações produzem a ativação ou a repressão da transcrição, dependendo de quais aminoácidos na cauda da histona são metilados. Enzimas chamadas de histona metiltransferases adicionam grupos metila a aminoácidos específicos (em geral, lisina ou arginina) das proteínas histona. Outras enzimas, chamadas de demetilases, removem os grupos metila das histonas. Muitas das enzimas e proteínas que modificam as histonas, como as metiltransferases e as demetilases, não se ligam a

sequências de DNA específicas, e devem ser recrutadas para sítios específicos da cromatina. Proteínas ligadoras de sequências específicas, modificações de histona preexistente e moléculas de RNA recrutam enzimas modificadoras de histona para sítios específicos.

Uma modificação comum é a adição de três grupos metila à lisina 4 na cauda da proteína histona H3, abreviada como H3K4me3 (K é a abreviação para lisina). As histonas com essa modificação são frequentemente encontradas próximo a promotores de genes ativos para transcrição nos eucariotos. Estudos identificaram proteínas que reconhecem e se ligam à H3K4me3, incluindo o fator de remodelagem de nucleossomo (NURF). O NURF e outras proteínas que reconhecem a H3K4me3 têm um domínio de ligação à proteína comum que se liga à cauda da histona H3 e, então, altera o empacotamento da cromatina, possibilitando a transcrição. A pesquisa também demonstrou que alguns fatores de transcrição, necessários para o início da transcrição (ver *Capítulo 13* e *Seção 17.3*), ligam-se diretamente à H3K4me3.

Acetilação das histonas. Outro tipo de modificação da histona que afeta a estrutura da cromatina é a acetilação, a adição de grupos acetila (CH_3CO) à histona. Em geral, a acetilação das histonas estimula a transcrição. Por exemplo, a adição de um único grupo acetila à lisina 16 na cauda da histona H4 impede a formação da fibra de cromatina de 30 nm (ver Figura 11.3), fazendo com que a cromatina esteja em uma configuração aberta e disponível para transcrição. Em geral, os grupos acetila desestabilizam a estrutura da cromatina, possibilitando a transcrição (**Figura 17.2**). Grupos acetila são adicionados a proteínas histona pelas enzimas acetiltransferase; outras enzimas, chamadas de desacetilases, desnudam os grupos acetila das histonas e restauram a estrutura da cromatina, que reprime a transcrição. Alguns fatores de transcrição (ver *Capítulo 13*) e outras proteínas que também regulam a transcrição têm atividade acetiltransferase ou atraem acetiltransferases ao DNA.

A acetilação das histonas controla a floração na Arabidopsis. A importância da acetilação da histona na regulação gênica é demonstrada pelo controle da floração na *Arabidopsis thaliana*, uma planta com várias características que a tornam um excelente modelo genético para sistemas das plantas (consulte mais informações no *Guia de Referência para Organismos-Modelo da Genética*, no final deste livro). O momento no qual a floração ocorre é crítico para a vida de uma planta; se a floração for iniciada na época errada no ano, os polinizadores podem não estar disponíveis para fertilizar as flores ou as condições ambientais podem não ser adequadas para a sobrevivência e a germinação das sementes. Consequentemente, o momento da floração na maioria das plantas é cuidadosamente regulado em resposta a múltiplas indicações internas e externas, como tamanho da planta, fotoperíodo e temperatura.

Entre os muitos genes que controlam a floração na *Arabidopsis* está o *locus C de floração* (*FLC*), que participa na supressão da floração após um período prolongado de frio (um processo chamado de vernalização). O gene *FLC* codifica uma proteína reguladora que reprime a atividade de outros genes que afetam a floração (**Figura 17.3**). Enquanto o *FLC* estiver ativo, a floração permanece suprimida. A atividade do *FLC* é controlada por outro *locus*, chamado *locus D de floração* (*FLD*), cuja função-chave é estimular a floração ao reprimir a ação do *FLC*. Em essência, a floração é estimulada porque o *FLD* reprime o repressor. Como *FLD* reprime *FLC*?

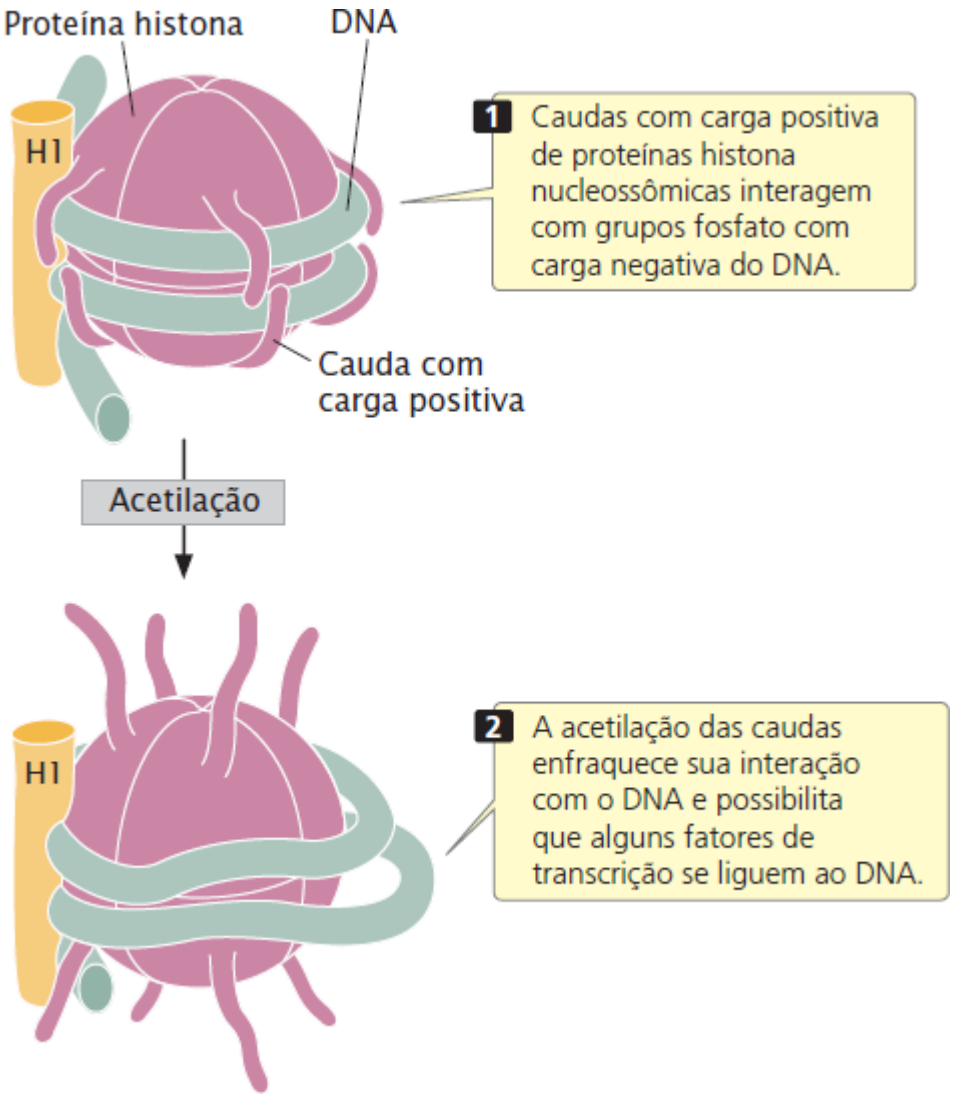


Figura 17.2 A acetilação das proteínas histona altera a estrutura da cromatina e possibilita que alguns fatores de transcrição se liguem ao DNA.

O *FLD* codifica uma enzima desacetilase, que remove os grupos acetila das proteínas histona na cromatina que circunda o *FLC* (ver Figura 17.3). A remoção dos grupos acetila das histonas altera a estrutura da cromatina e inibe a transcrição. A inibição da transcrição impede que o *FLC* seja transcrito e remove sua repressão sobre a floração. Resumindo, o *FLD* estimula a floração na *Arabidopsis* ao desacetilar a cromatina que circunda o *FLC*, removendo seu efeito inibidor sobre a floração.

As mudanças na estrutura da cromatina que afetam a expressão gênica, como os mecanismos já descritos (remodelagem da cromatina, modificação da histona e metilação do DNA), são exemplos do fenômeno da epigenética. A epigenética é definida como alterações no DNA e na estrutura da cromatina que afetam traços e são transmitidas para outras células ou gerações, mas não são causadas por mudanças nas sequências de bases do DNA. O Capítulo 21 descreve as mudanças epigenéticas e seus efeitos sobre a regulação gênica com mais detalhes. Resolva o Problema 18

Imunoprecipitação da cromatina. Nossa compreensão sobre como as mudanças na estrutura da cromatina estão associadas à expressão gênica, ou como as proteínas ligadoras do DNA afetam a transcrição, avançaram muito com o uso de uma técnica chamada imunoprecipitação da cromatina (ChIP). Essa técnica torna possível que os pesquisadores determinem localizações específicas dentro do genoma onde as proteínas interagem com o DNA. Essas proteínas podem ser histonas que sofreram modificações, fatores de transcrição ou outras proteínas que se ligam a promotores e acentuadores (sequências que afetam a transcrição de genes distantes, ver *Capítulo 13*).

A ideia básica da ChIP é que uma proteína particular e o DNA ao qual ela está ligada são isolados, a proteína e o DNA são separados e a sequência de DNA ligada à proteína é identificada. A técnica fornece meios poderosos para determinar as localizações amplas de genoma de histonas modificadas e os sítios de ligação para os fatores de transcrição e outras proteínas que afetam a transcrição.

Uma versão da ChIP, chamada ChIP de ligação cruzada (XChIP), é usada para identificar os sítios de ligação dos fatores de transcrição e de outras proteínas que se ligam ao DNA. Nesse procedimento (**Figura 17.4**), a proteína e o DNA associado apresentam ligação cruzada temporária, o que significa que eles são tratados com formaldeído ou luz UV para criar ligações

covalentes entre o DNA e a proteína. As ligações cruzadas mantêm o DNA e a proteína unidos de modo que esse DNA será separado com a proteína. Após a ligação cruzada, a célula é rompida e a cromatina é quebrada em pedaços pela digestão de uma enzima ou por corte mecânico. São aplicados anticorpos específicos para uma determinada proteína – como um fator de transcrição específico. Os anticorpos se fixam aos complexos proteína-DNA e provocam sua precipitação. A ligação cruzada é revertida, separando o DNA da proteína. A proteína é removida por uma enzima que a digere, mas que não digere o DNA, deixando fragmentos do DNA ao qual a proteína estava ligada. Esses fragmentos podem ser sequenciados ou identificados com outros métodos. O resultado é a informação sobre as localizações genômicas dos sítios de ligação para a proteína específica.

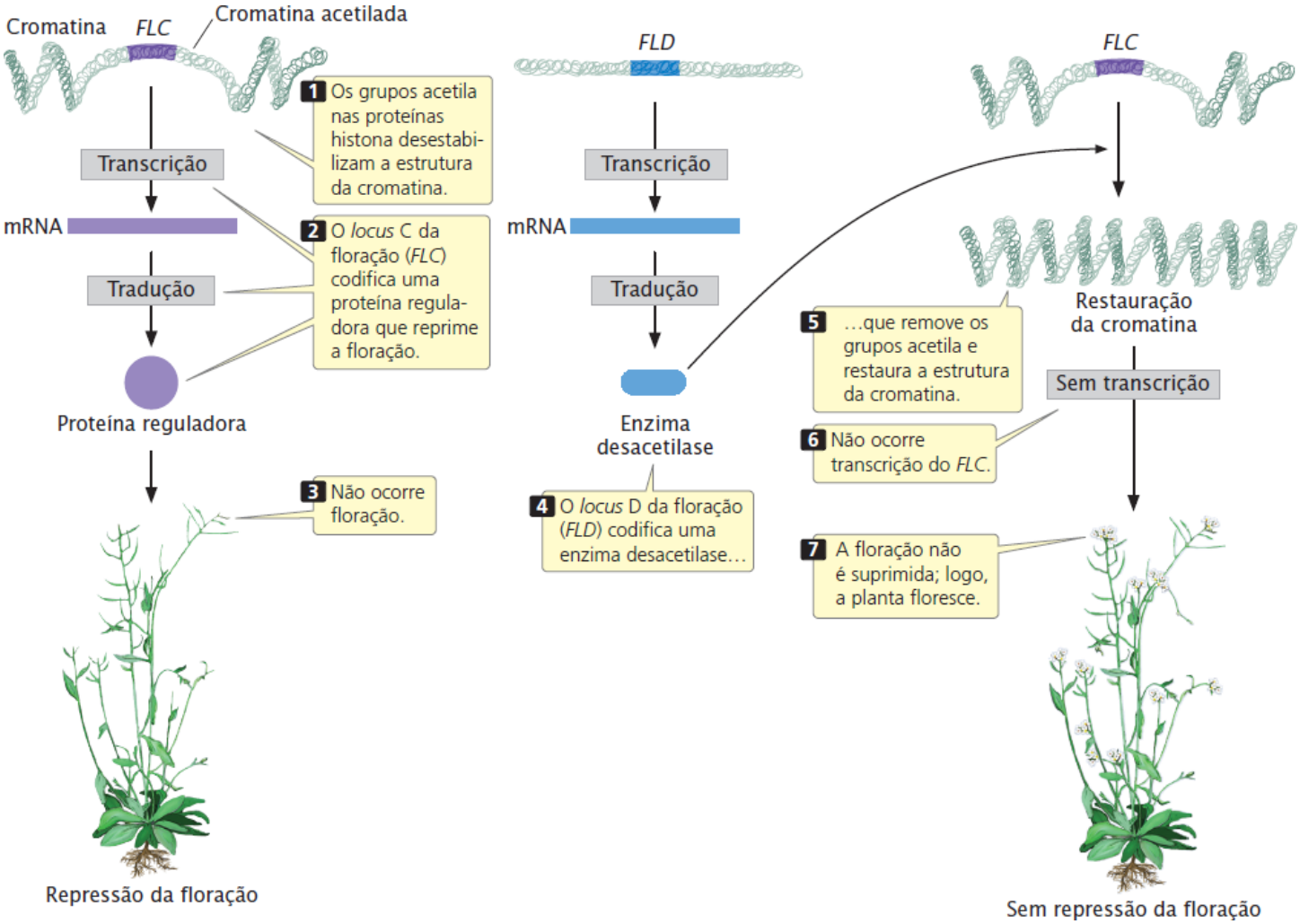


Figura 17.3 A floração na *Arabidopsis* é controlada em parte pelo *FLD*, um gene que codifica uma enzima desacetilase. Essa enzima remove os grupos acetila das proteínas histona na cromatina que circunda o *FLC*, um gene que suprime a floração. A remoção dos grupos acetila das histonas restaura a estrutura da cromatina e reprime a transcrição do *FLC*, deixando a planta florescer.

Uma versão dessa técnica, chamada ChIP nativa (nChIP), não usa ligação cruzada. Ela é empregada para encontrar as localizações das proteínas histona modificadas. Nesse caso, não é necessária a ligação cruzada porque o DNA e as proteínas histona estão naturalmente ligados pela estrutura do nucleossomo. A cromatina é isolada da célula e fragmentada, e os anticorpos para uma proteína particular (em geral uma histona modificada específica) são usados para precipitar os complexos proteína-DNA. A proteína e o DNA são separados; a proteína é digerida e os fragmentos de DNA aos quais as histonas modificadas estão presas são sequenciados ou identificados de outra maneira.

A análise por ChIP tem sido considerada para determinar as localizações de histonas modificadas que ativam ou reprimem a transcrição. Como já foi mencionado, a marca da histona H3K4me3 está associada a promotores de genes ativos. A análise nChIP identifica com sucesso as localizações dessa histona modificada no genoma humano, ajudando a identificar os promotores ativos em diferentes tecidos.

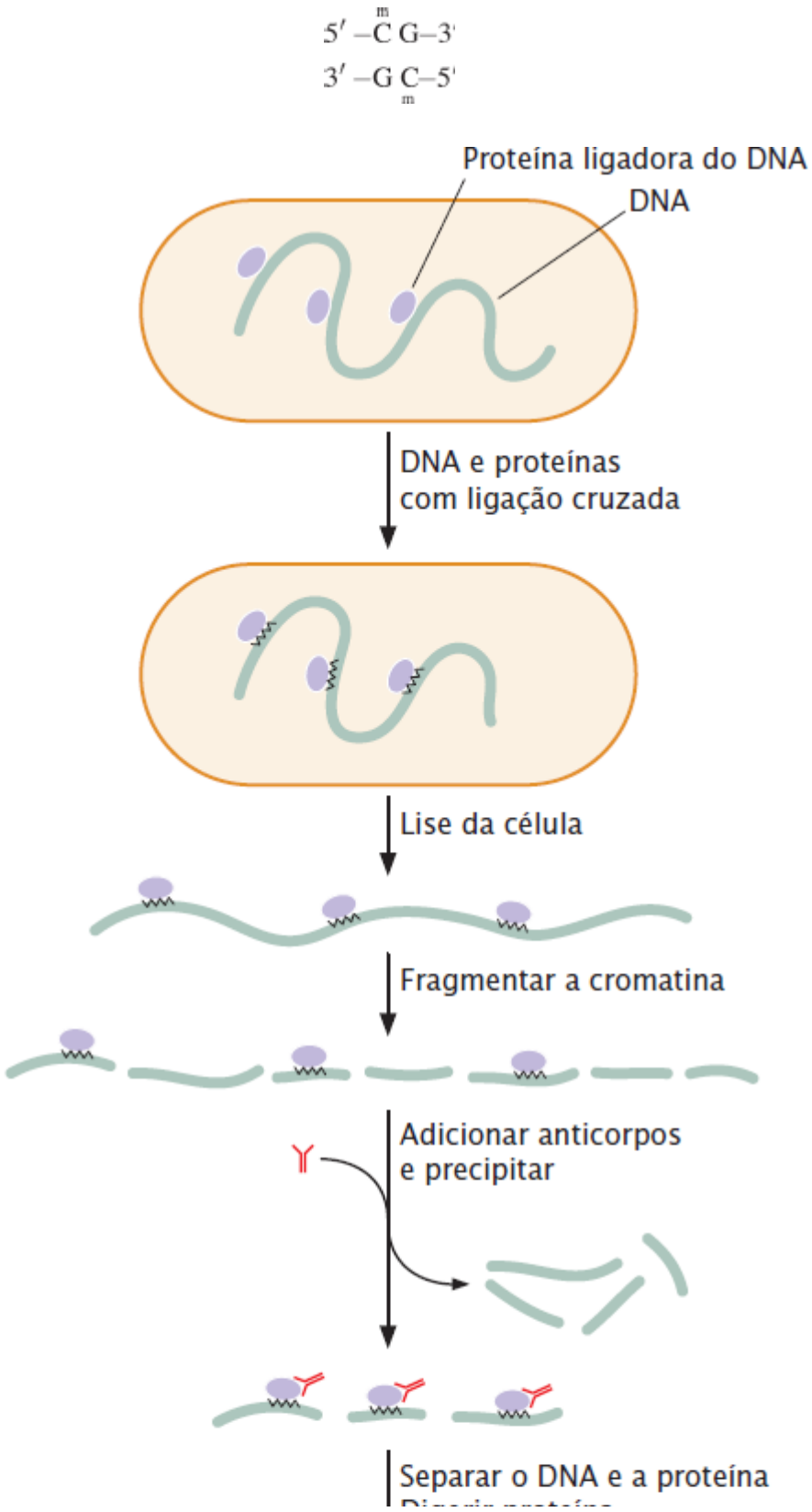
Conceitos

As caudas das proteínas histonas são modificadas por adição ou remoção de grupos fosfato, metila ou acetila ou por outras alterações. Essas modificações alteram a estrutura da cromatina e afetam a transcrição dos genes.

Metilação do DNA

Outra mudança na estrutura da cromatina associada à transcrição é a metilação das bases de citosina, que geram a 5-metilcitosina (ver Figura 10.19). A metilação da citosina no DNA é diferente da metilação das proteínas histona já discutida. O DNA muito metilado está associado à repressão da transcrição nos vertebrados e nas plantas, enquanto o DNA ativo para transcrição em geral não está metilado nesses organismos. Padrões anormais de metilação também estão associados a alguns tipos de câncer.

A metilação do DNA é mais comum nas bases citosina adjacentes aos nucleotídios guanina (CpG, em que p representa o grupo fosfato no esqueleto do DNA); então, duas citosinas metiladas encontram-se em diagonal cruzada uma da outra nas fitas opostas:



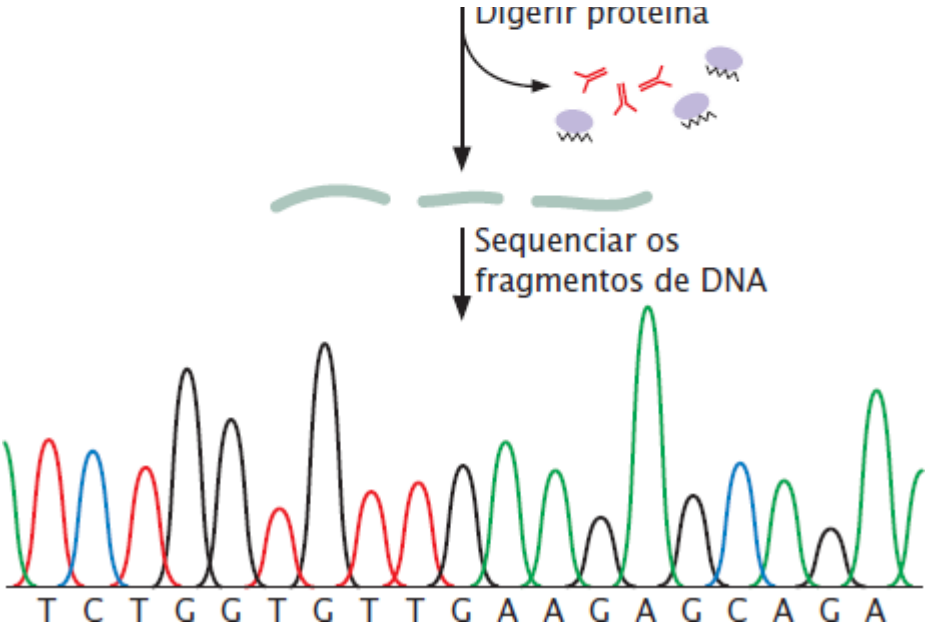


Figura 17.4 A imunoprecipitação da cromatina (ChIP) pode ser usada para identificar sítios de ligação ao DNA de uma proteína específica e as localizações de proteínas histona modificadas. Este esquema mostra o método usado para a imunoprecipitação de cromatina por ligação cruzada (XChIP).

As regiões do DNA com muitas sequências CpG são chamadas de **ilhas CpG** e são comumente encontradas próximo dos sítios de início de transcrição. Enquanto os genes não estão sendo transcritos, essas ilhas CpG estão metiladas, mas os grupos metila são removidos antes do início da transcrição. A metilação CpG também está associada à repressão gênica a longo prazo, como no cromossomo X inativado dos mamíferos fêmeas (ver *Capítulo 4*).

Evidências indicam que existe uma associação entre a metilação do DNA e a desacetilação das histonas, ambas reprimindo a transcrição. Algumas proteínas que se ligam firmemente a sequências CpG metiladas formam complexos com outras proteínas que atuam como histona desacetilases. Em outras palavras, a metilação atrai as desacetilases, que removem os grupos acetila a partir das caudas de histona, estabilizando a estrutura do nucleossomo e reprimindo a transcrição. A desmetilação do DNA possibilita que as acetiltransferases adicionem grupos acetila, rompendo a estrutura do nucleossomo e promovendo a transcrição. Resolva o Problema 19

Conceitos

A estrutura da cromatina pode ser alterada pela metilação do DNA. Nos eucariotos, a 5-metilcitosina encontrada nos dinucleotídios CpG é o componente da metilação do DNA. A metilação do DNA está associada à repressão da transcrição.

Checagem dos conceitos 1

Cite alguns dos processos diferentes que afetam a regulação gênica ao alterar a estrutura da cromatina.

17.3 O início da transcrição é regulado por fatores de transcrição e proteínas reguladoras

Analisamos apenas um nível no qual a expressão gênica é controlada – a alteração da cromatina e da estrutura do DNA. Agora acompanharemos outro nível importante de controle – feito por meio da ligação de proteínas a sequências de DNA que afetam a transcrição. A transcrição é um nível importante de controle nas células eucarióticas, e esse controle demanda alguns tipos diferentes de proteínas e elementos regulatórios. O início da transcrição eucariótica foi discutido em detalhes no Capítulo 13. Lembre-se de que os fatores gerais de transcrição e a RNA polimerase se montam em um *aparato basal de transcrição*, o complexo de RNA polimerase, os fatores de transcrição e outras proteínas responsáveis pela transcrição. O aparato basal de transcrição se liga ao promotor central localizado imediatamente *upstream* de um gene e é capaz de ativar níveis mínimos de transcrição; as *proteínas reguladoras da transcrição* são necessárias para promover a transcrição em níveis normais. Essas proteínas se ligam a um promotor regulatório, localizado *upstream* do promotor central (**Figura 17.5**), e aos *acentuadores*, que estão localizados a alguma distância do gene. Algumas proteínas reguladoras da transcrição são ativadoras, estimulando a transcrição. Outras são repressoras, inibindo a transcrição.

Ativadores e coativadores da transcrição

As proteínas ativadoras da transcrição estimulam e estabilizam o aparato basal de transcrição no promotor central. Os ativadores interagem diretamente com o aparato basal de transcrição ou indiretamente, pelas proteínas coativadoras. Alguns ativadores e coativadores, assim como os fatores gerais de transcrição, também têm atividade acetiltransferase e estimulam a transcrição ao alterar a estrutura da cromatina.

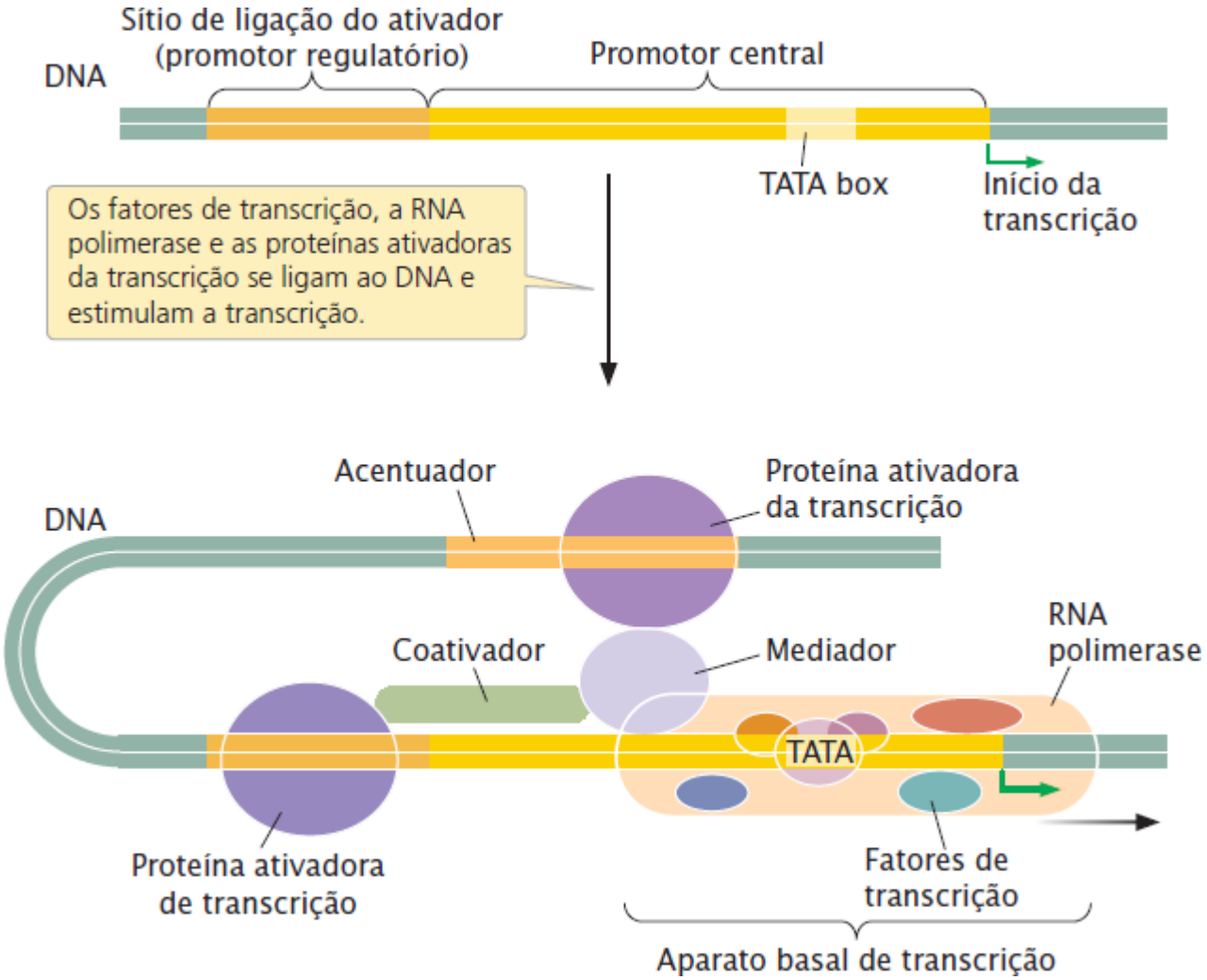


Figura 17.6 As sequências consenso nos promotores dos três genes eucarióticos ilustram o princípio de que diferentes sequências podem ser misturadas e unidas em diferentes combinações. Uma proteína ativadora da transcrição diferente se liga a cada sequência consenso, então cada promotor responde a uma combinação única de proteínas ativadoras.

As proteínas ativadoras da transcrição têm duas funções distintas (ver Figura 17.5). Primeiro, elas são capazes de se ligar ao DNA em uma sequência de base específica, em geral uma sequência consenso em um promotor regulatório ou acentuador; para essa função, a maioria das proteínas ativadoras da transcrição tem um ou mais motivos de ligação ao DNA, como hélice-giro-hélice, dedo-de-zinco e zíper de leucina (ver *Capítulo 16*). Uma segunda função é a capacidade de interagir com outros componentes do aparato de transcrição e influenciar a taxa de transcrição.

Dentro do promotor regulatório há várias sequências consenso distintas, às quais diferentes ativadores da transcrição se ligam. Entre promotores distintos, os sítios de ligação do ativador são misturados e unidos em diferentes combinações (**Figura 17.6**) e cada promotor é regulado por uma combinação única das proteínas ativadoras da transcrição.

As proteínas ativadoras da transcrição se ligam a sequências no promotor regulatório e afetam a montagem ou a estabilidade do aparato basal de transcrição no promotor central. Um dos componentes do aparato basal de transcrição é um complexo de proteínas chamado de **mediador** (ver Figura 17.6). As proteínas ativadoras da transcrição que se ligam a sequências no promotor regulatório (ou acentuador, ver a próxima seção) fazem contato com o mediador e afetam a taxa na qual a transcrição é iniciada. Alguns promotores regulatórios também têm sequências ligadas a proteínas que reduzem a taxa de transcrição por meio de interações inibitórias com o mediador.

Regulação do metabolismo da galactose por meio de GAL4. Um exemplo de proteína ativadora da transcrição é a

GAL4, que regula a transcrição de vários genes da levedura cujos produtos metabolizam a galactose. Como os genes no óperon *lac*, os genes que controlam o metabolismo da galactose são induzíveis: quando não há galactose, esses genes não são transcritos e as proteínas que degradam a galactose não são produzidas; quando há galactose, os genes são transcritos e as enzimas são sintetizadas. A GAL4 tem vários dedos de zinco e se liga a uma sequência de DNA chamada UAS_G (sequência ativadora *upstream* para GAL4). A UAS_G exibe as propriedades de um acentuador – uma sequência regulatória que pode estar distante do gene regulado e é independente do gene em posição e orientação (ver *Capítulo 13*). Quando ligada à UAS_G, a GAL4 ativa a transcrição dos genes da levedura necessários para metabolizar a galactose. A GAL4 e várias outras proteínas ativadoras da transcrição têm múltiplos aminoácidos com cargas negativas que formam um *domínio de ativação ácido*. Esses ativadores ácidos estimulam a transcrição ao estimular a montagem do aparato basal de transcrição.

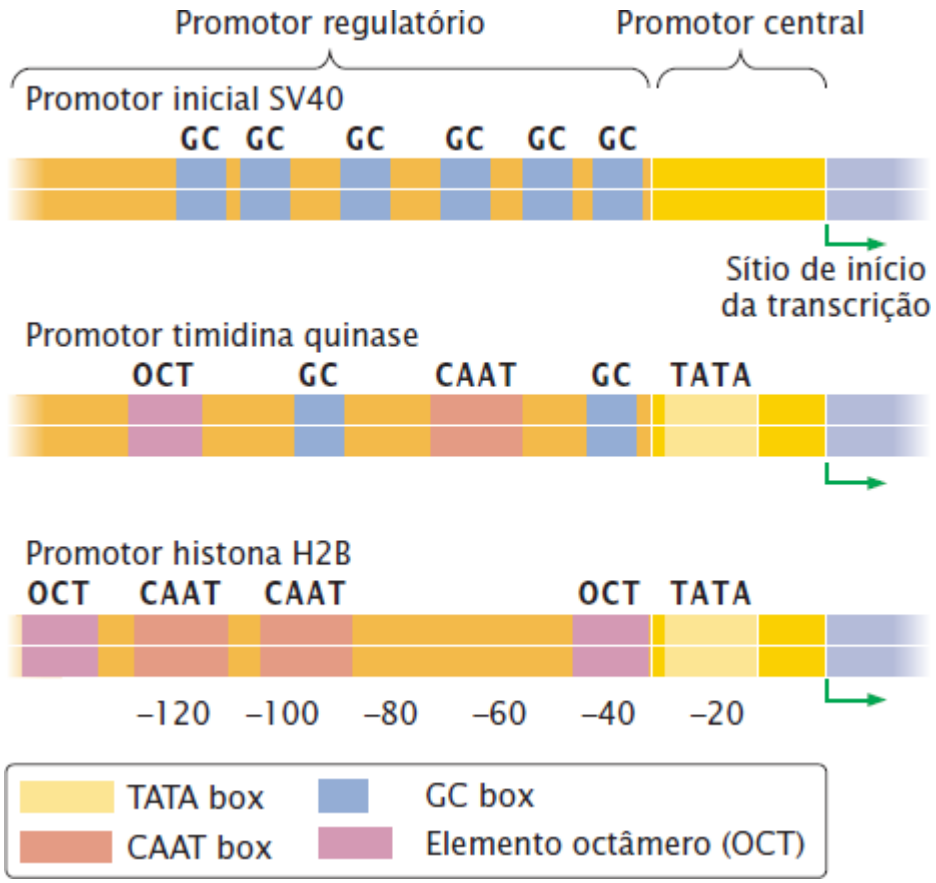


Figura 17.5 As proteínas ativadoras da transcrição se ligam a sítios no DNA e estimulam a transcrição. A maioria age ao estimular ou estabilizar a montagem do aparato basal de transcrição.

Uma região específica da GAL4 se liga a outra proteína chamada GAL80, que regula a atividade da GAL4 quando há galactose. Quando não há galactose, a GAL80 se liga à GAL4, impedindo que a GAL4 ative a transcrição (**Figura 17.7**). Quando há galactose, entretanto, ela se liga a outra proteína chamada GAL3, que interage com a GAL80, provocando uma mudança conformacional na GAL80 de modo que ela não pode mais se ligar à GAL4. A proteína GAL4 fica livre para ativar a transcrição dos genes, cujos produtos metabolizam a galactose.

Repressores da transcrição

Algumas proteínas regulatórias nas células eucarióticas atuam como repressores, inibindo a transcrição. Esses repressores se ligam a sequências no promotor regulatório ou a sequências distantes, chamadas de *silenciadores*, que, como os acentuadores, são independentes em relação à posição e orientação. Diferente dos repressores nas bactérias, a maioria dos repressores eucarióticos não bloqueia diretamente a RNA polimerase. Esses repressores podem competir com os ativadores pelos sítios de ligação ao DNA: quando um sítio é ocupado por um ativador, a transcrição é ativada, mas não há ativação se um repressor ocupar esse sítio. Por outro lado, um repressor pode se ligar a sítios próximos de um sítio ativador e evitar que o ativador entre em contato com o aparato basal de transcrição. Um terceiro mecanismo possível de ação do repressor é a interferência direta com a montagem do aparato basal de transcrição, bloqueando o início da transcrição.

Conceitos

As proteínas regulatórias da transcrição nas células eucarióticas podem influenciar o início da transcrição ao afetar a estabilidade ou a montagem do aparato basal de transcrição. Algumas proteínas regulatórias são ativadoras e estimulam a transcrição; outras são repressoras e inibem o início da transcrição.

Checagem dos conceitos 2

A maioria das proteínas ativadoras da transcrição interagem com:

a. *Íntrons*.
b. O aparato basal de transcrição.
c. A DNA polimerase.
d. O terminador.

Acentuadores e insuladores

Os acentuadores são capazes de afetar a transcrição em promotores distantes. Por exemplo, um acentuador que regula o gene que codifica a cadeia alfa do receptor do linfócito T está a 69.000 pb *downstream* do promotor do gene. Além disso, a posição exata e a orientação de um acentuador relativo ao promotor podem variar. Como um acentuador pode afetar o início da transcrição que ocorre em um promotor a 10.000 pb de distância? Em muitos casos, as proteínas reguladoras se ligam ao acentuador e fazem com que o DNA entre o acentuador e o promotor se afrouxe, aproximando o promotor e o acentuador, e então as proteínas reguladoras da transcrição são capazes de interagir diretamente com o aparato basal de transcrição no promotor central (ver Figura 17.5). Alguns acentuadores podem ser atraídos aos promotores por proteínas que se ligam a sequências no promotor regulatório e "amarram" o acentuador próximo do promotor central. Os acentuadores também podem afetar a transcrição por meio de modificações que alteram a estrutura da cromatina. Um acentuador típico tem 500 pb de comprimento e 10 sítios de ligação para proteínas que regulam a transcrição.

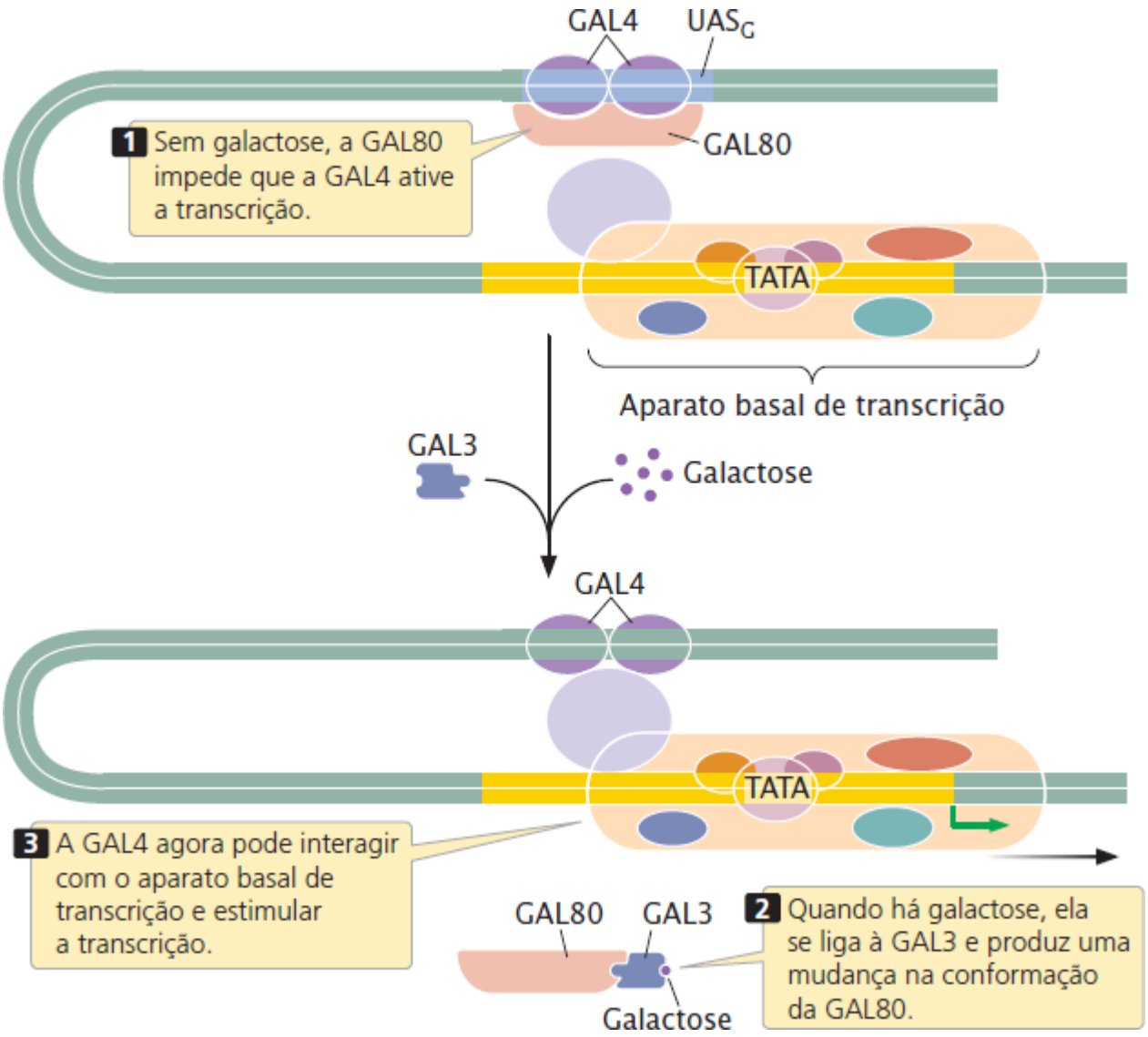


Figura 17.7 A transcrição é ativada pela GAL4 em resposta à galactose. A GAL4 se liga ao sítio da UAS_G e controla a transcrição dos genes no metabolismo da galactose.

Pesquisa recente demonstrou que muitos acentuadores por si sós são transcritos em pequenas moléculas de RNA chamadas de RNAs acentuadores (eRNAs). A evidência sugere que a transcrição de um acentuador esteja associada à transcrição nos promotores que os acentuadores afetam. Mas não está claro como a transcrição no acentuador pode afetar a transcrição em um promotor distante. Os acentuadores podem recrutar a RNA polimerase, que é então transferida para o promotor quando o acentuador interage com o promotor. Por outro lado, a transcrição do acentuador possibilita que a cromatina adote uma configuração mais aberta, o que facilita a transcrição nos promotores próximos.

A maioria dos acentuadores é capaz de estimular qualquer promotor na sua vizinhança. Seus efeitos são limitados, entretanto, pelos **insuladores** (também chamados de elementos de fronteira), sequências de DNA que bloqueiam ou isolam o efeito dos acentuadores, dependendo da posição. Se o insulador estiver entre o acentuador e o promotor, ele bloqueia a ação do acentuador; mas o insulador não tem efeito se estiver fora da região entre eles (**Figura 17.8**). Proteínas específicas se ligam aos insuladores e participam na sua atividade bloqueadora. Alguns insuladores também limitam a difusão das mudanças na estrutura da cromatina que afetam a transcrição. Os procariotos têm alguns elementos semelhantes aos acentuadores. **Resolva o Problema 23**

Conceitos

Algumas proteínas regulatórias se ligam aos acentuadores, que são elementos regulatórios distantes do gene cuja transcrição estimulam. Os insuladores são sequências de DNA que bloqueiam a ação dos acentuadores.

Checagem dos conceitos 3

Como a ligação das proteínas regulatórias aos acentuadores afeta a transcrição em genes que estão distantes milhares de pares de base?

Regulação da parada e alongamento transcricional

A transcrição nos eucariotos é regulada por fatores que afetam o início da transcrição, incluindo mudanças na estrutura da cromatina, fatores de transcrição e proteínas regulatórias de transcrição. A pesquisa indica que a transcrição também pode ser controlada por fatores que afetam a parada e o alongamento da RNA polimerase após o início da transcrição.

O aparato basal de transcrição – composto por RNA polimerase, fatores de transcrição e outras proteínas – monta-se no promotor central. Quando ocorre o início da transcrição, a RNA polimerase se move *downstream*, transcrevendo o gene estrutural e produzindo um RNA. Em alguns genes, a RNA polimerase inicia a transcrição e transcreve de 24 a 50 nucleotídios de RNA, mas então pausa ou para. Por exemplo, a parada é observada em genes que codificam **proteínas de choque térmico** na *Drosophila* – proteínas que ajudam a evitar o dano por agentes prejudiciais, como o calor extremo. As proteínas de choque térmico são produzidas por vários genes diferentes. Durante os períodos de estresse ambiental, a transcrição de todos os genes de choque térmico está muito elevada. A RNA polimerase inicia a transcrição nos genes de choque térmico na *Drosophila*, mas, na ausência de estresse, para *downstream* do sítio de início de transcrição. As polimerases paradas são liberadas quando é encontrado um fator de estresse, possibilitando a transcrição rápida dos genes e a produção de proteínas de choque térmico que facilitam a adaptação ao ambiente estressante.

Acreditava-se que a parada ocorresse apenas em um pequeno número de genes, mas a pesquisa atual indica que ela é disseminada por todo o genoma eucariótico. Por exemplo, RNAs polimerases paradas foram encontradas em centenas de genes na *Drosophila*. Foram identificados vários fatores que promovem a parada; um deles é uma proteína chamada fator de alongamento negativo (NELF), que se liga à RNA polimerase e a paralisa após o início da transcrição. Outra proteína, chamada fator de alongamento da transcrição b positivo (P-TEFb), libera a parada e promove o alongamento por meio da fosforilação do NELF e da RNA polimerase, talvez ao fazer o NELF se dissociar da polimerase.

Conceitos

Em alguns genes, a RNA polimerase pode pausar ou parar *downstream* do promotor. Os fatores regulatórios afetam a parada e o alongamento da transcrição.

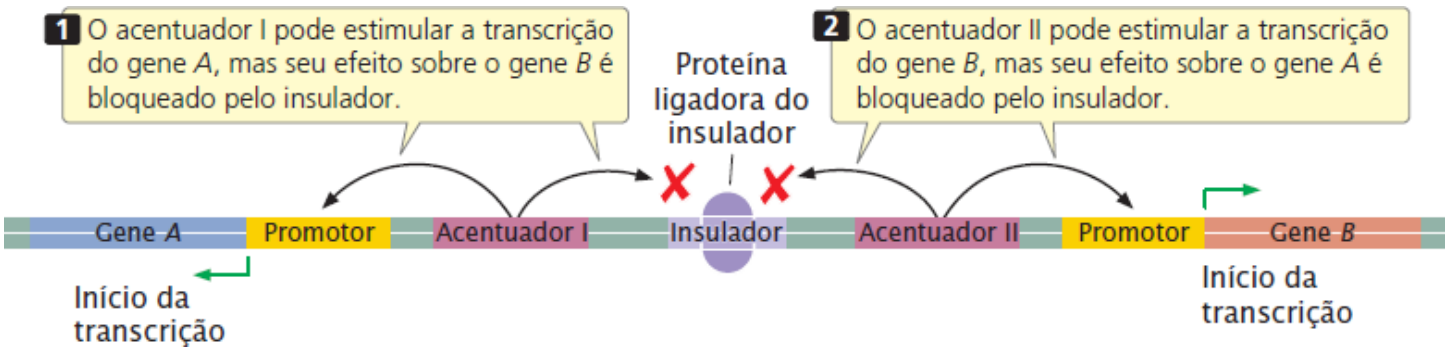


Figura 17.8 **Um insulador bloqueia a ação de um acentuador em um promotor quando o insulador está entre o acentuador e o promotor.**

Regulação gênica coordenada

Embora a maioria das células eucarióticas não tenha óperons, vários genes eucarióticos podem ser ativados pelo mesmo estímulo. Grupos de genes bacterianos são expressos de forma coordenada (ligados e desligados juntos) porque estão fisicamente agrupados como um óperon e têm o mesmo promotor, mas os genes expressos de forma coordenada nas células eucarióticas não estão agrupados. Como, então, a transcrição dos genes eucarióticos é controlada de forma coordenada se eles não estão organizados em um óperon?

Genes expressos de forma coordenada nas células eucarióticas são capazes de responder ao mesmo estímulo porque têm sequências regulatórias pequenas em comum com seus promotores ou acentuadores. Por exemplo, diferentes genes eucarióticos para choque término apresentam um elemento regulatório comum *upstream* dos seus sítios de início. Tais sequências regulatórias de DNA são chamadas de **elementos de resposta**; são sequências curtas que tipicamente têm sequências consenso (**Quadro 17.1**) em distâncias variadas a partir do gene a ser regulado. Os elementos de resposta são sítios de ligação para os ativadores da transcrição. Uma proteína ativadora da transcrição se liga ao elemento de resposta e eleva a transcrição. O mesmo elemento de resposta pode ser encontrado em diferentes genes, possibilitando que múltiplos genes sejam ativados pelo mesmo estímulo.

Um único gene eucariótico pode ser regulado por vários elementos de resposta diferentes. Por exemplo, o gene metalotioneína protege as células da toxicidade dos metais pesados ao codificar uma proteína que se liga aos metais pesados e os remove das células. O aparato basal de transcrição se monta ao redor da TATA box, *upstream* do sítio de início de transcrição para o gene metalotioneína, mas o aparato sozinho é capaz de produzir apenas baixas taxas de transcrição.

Outros elementos de resposta encontrados *upstream* do gene metalotioneína contribuem para o aumento da sua taxa de transcrição. Por exemplo, várias cópias de um elemento de resposta a metal (MRE) estão *upstream* do gene da metalotioneína (**Figura 17.9**). Os metais pesados estimulam a ligação das proteínas ativadoras aos MREs, que eleva a taxa de transcrição do gene da metalotioneína. Como existem múltiplas cópias de MRE, as taxas elevadas de transcrição são induzidas pelos metais. Dois acentuadores também estão localizados na região *upstream* do gene metalotioneína; um deles tem um elemento de resposta conhecido como TRE, que estimula a transcrição quando há uma proteína ativada chamada AP1. Um terceiro elemento de resposta, chamado GRE, está localizado a aproximadamente 250 nucleotídios *upstream* do gene da metalotioneína e estimula a transcrição em resposta a alguns hormônios.

Esse exemplo ilustra uma característica comum do controle transcricional eucariótico: um único gene pode ser ativado por vários elementos de resposta diferentes encontrados nos promotores e acentuadores. Os múltiplos elementos de resposta levam o mesmo gene a ser ativado por diferentes estímulos. Ao mesmo tempo, o mesmo elemento de resposta em diferentes genes possibilita um único estímulo para ativar múltiplos genes. Desse modo, os elementos de resposta tornam possível respostas bioquímicas complexas nas células eucarióticas.

17.4 Alguns genes são regulados por processamento e degradação do RNA

Nas bactérias, a transcrição e a tradução ocorrem simultaneamente. Nos eucariotos, a transcrição acontece no núcleo e os pré-mRNAs são processados antes de se deslocarem para o citoplasma para tradução, criando oportunidades para controle gênico após a transcrição. Consequentemente, a regulação gênica pós-transcricional é importante nas células eucarióticas. Um nível comum da regulação gênica nos eucariotos é o processamento e a degradação do RNA.

Quadro 17.1 Alguns elementos de resposta encontrados nas células eucarióticas.

Elemento de resposta	Responde a	Sequência consenso
Elemento de choque térmico	Calor e outros estresses	CNNGAANNTCCNNG
Elemento de resposta ao glicocorticoide	Glicocorticoides	TGGTACAAATGTTCT
Elemento de resposta ao éster de forbol	Ésteres de forbol	TGACTCA
Elemento de resposta sérico	Soro	CCATATTAGG

Fonte: Segundo B. Lewin, *Genes IV* (Oxford University Press, 1994), p. 880.

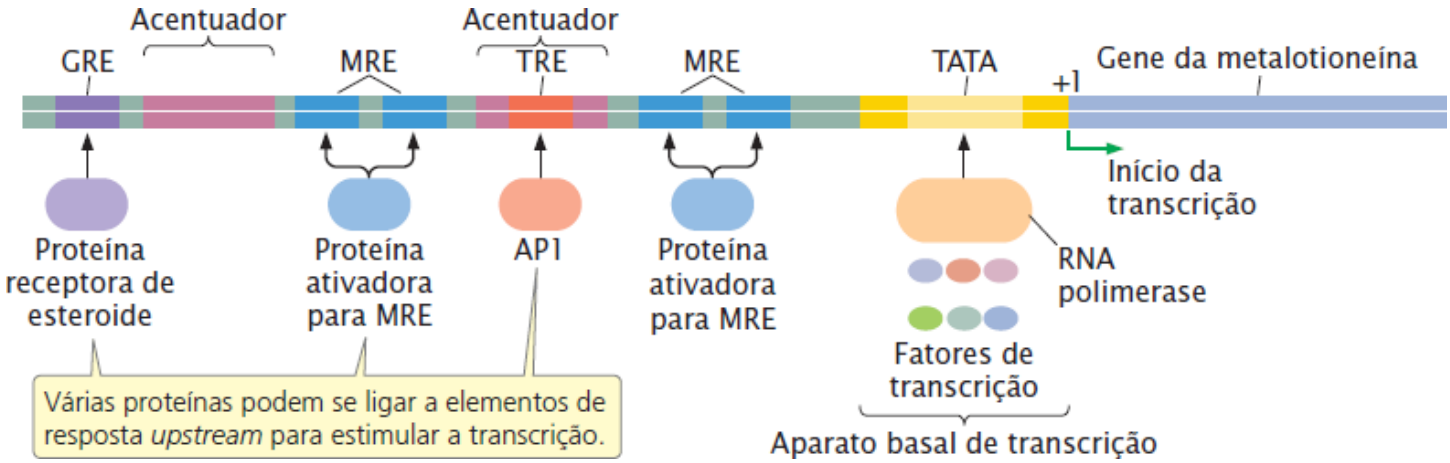


Figura 17.9 Elementos múltiplos de resposta (MREs) são encontrados na região *upstream* do gene da metalotioneína. O aparato basal de transcrição se liga próximo à TATA box. Em resposta aos metais pesados, as proteínas ativadoras se ligam a vários MREs e estimulam a transcrição. O elemento de resposta ao TRE é o sítio de ligação para o fator de transcrição AP1, que é estimulado pelos ésteres de forbol. Em resposta aos hormônios glicocorticoides, as proteínas receptoras de esteroide se ligam ao elemento de resposta GRE, localizado a aproximadamente 250 nucleotídios *upstream* do gene da metalotioneína, e estimulam a transcrição.

Regulação gênica por meio da recomposição do RNA

A recomposição (*splicing*) alternativa possibilita que um pré-mRNA seja unido de várias maneiras, gerando diferentes proteínas em diferentes tecidos ou em distintos períodos do desenvolvimento (ver *Capítulo 14*). Muitos genes eucarióticos sofrem a recomposição alternativa, e a regulação da recomposição é um meio importante do controle da expressão gênica nas células eucarióticas.

Recomposição alternativa no gene do antígeno T. O gene do antígeno T do vírus SV40 nos mamíferos é um exemplo bem estudado de recomposição alternativa. Esse gene é capaz de codificar duas proteínas diferentes, os antígenos T grande e t pequeno. Qual das duas proteínas é produzida depende de qual dos dois sítios alternativos de recomposição 5′ é usado na recomposição do RNA (**Figura 17.10**). O uso de um sítio de recomposição 5′ produz mRNA que codifica o antígeno T grande, enquanto o uso de outro sítio de recomposição 5′ (que está muito *upstream*) produz um mRNA que codifica o antígeno t pequeno.

Uma proteína chamada fator de recomposição 2 (SF2) acentua a produção do mRNA, codificando o antígeno t pequeno (ver Figura 17.10). O fator de recomposição 2 tem dois domínios de ligação: um deles é uma região ligante de RNA, e o outro tem os aminoácidos serina e arginina alternados. Esses dois domínios são típicos de **proteínas SR** (ricas em serina e arginina), que participam da regulação da recomposição. O SF2 estimula a ligação das ribonucleoproteínas nucleares pequenas (snRNPs) para o sítio de recomposição 5′, uma das primeiras etapas na recomposição do RNA (ver *Capítulo 14*). O mecanismo preciso pelo qual as proteínas SR influenciam a escolha dos sítios de corte ainda é pouco compreendido. Um modelo sugere que as proteínas SR se liguem a sítios de recomposição específicos no mRNA e estimulem a fixação dos snRNPs, que então levam à recomposição do sítio.

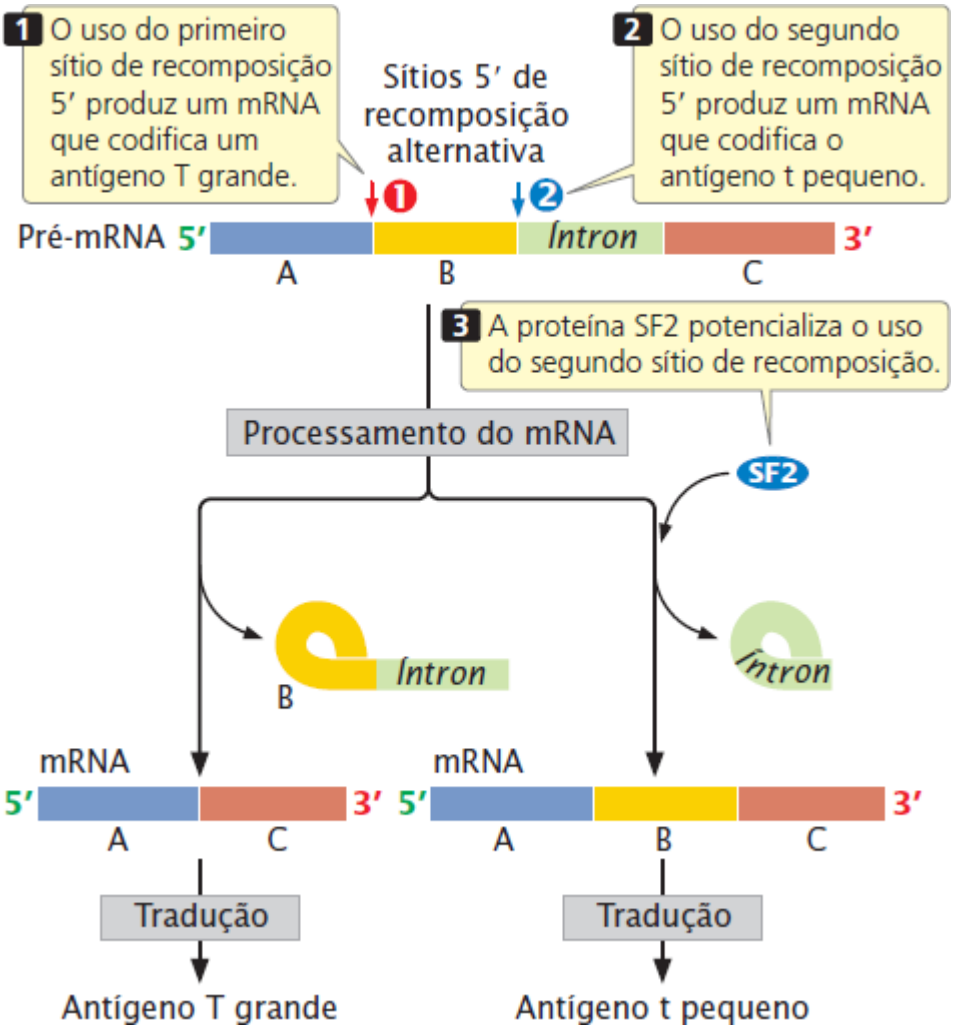


Figura 17.10 A recomposição alternativa leva à produção do antígeno t pequeno e do antígeno T grande no vírus SV40 de mamíferos.

Recomposição alternativa no desenvolvimento sexual de Drosophila. Outro exemplo de recomposição alternativa de mRNA que regula a expressão gênica é a determinação se uma mosca-da-fruta será macho ou fêmea. A diferenciação sexual na *Drosophila* surge a partir de uma cascata de regulação gênica. Quando existem dois cromossomos X, um promotor específico de fêmea é ativado no início do desenvolvimento e estimula a transcrição do gene *sexual letal* (*Sxl*) (**Figura 17.11).** A proteína codificada pelo *Sxl* regula a recomposição do pré-mRNA transcrito de outro gene chamado *transformador* (*tra*). A recomposição do pré-mRNA *tra* leva à produção da proteína Tra (ver Figura 17.11). Junto com outra proteína (Tra-2), a Tra estimula a recomposição específica de fêmea do pré-mRNA de outro gene chamado *sexo duplo* (*dsx*). Esse evento produz uma proteína Dsx específica para fêmea, que faz o embrião desenvolver características femininas.

Nos embriões machos, que têm um cromossomo X (ver Figura 17.11), o promotor que transcreve o gene *Sxl* nas fêmeas está inativo, e não há produção da proteína Sxl. Não havendo a proteína Sxl, o pré-mRNA *tra* é cortado em um local de recomposição 3′ diferente para produzir uma forma não funcional da proteína Tra (**Figura 17.12**). Por outro lado, esse Tra não funcional nos machos faz com que os pré-mRNAs *dsx* sejam unidos ou recompostos de maneira diferente do que nas fêmeas, levando à produção de uma proteína Dsx específica para macho (ver Figura 17.11). Esse evento leva ao desenvolvimento dos traços específicos dos machos.

Em resumo, as proteínas Tra, Tra-2 e Sxl regulam a recomposição alternativa que produz os fenótipos masculino e feminino na *Drosophila*. Ainda não se sabe exatamente como essas proteínas regulam a recomposição alternativa, mas é possível que a proteína Sxl (produzida apenas nas fêmeas) bloqueie o sítio de recomposição *upstream* do pré-mRNA *tra*. Esse bloqueio forçaria o spliceossomo a usar o sítio de recomposição 3′ *downstream*, o que levaria à produção da proteína Tra e acabaria resultando no fenótipo feminino (ver Figura 17.12). Resolva o Problema 24

Conceitos

Os genes eucarióticos podem ser regulados por controle do processamento do mRNA. A seleção de sítios de recomposição alternativa leva à produção de diferentes proteínas.

Degradação do RNA

A quantidade sintetizada de uma proteína depende da quantidade de mRNA correspondente disponível para tradução. A quantidade de mRNA disponível, por outro lado, depende tanto da taxa de síntese de mRNA quanto da taxa de degradação do mRNA. Os mRNAs eucarióticos são mais estáveis que os mRNAs bacterianos, que tipicamente duram apenas alguns minutos antes de serem degradados. Contudo, existe uma grande variação na estabilidade do mRNA eucariótico: alguns mRNAs duram apenas alguns minutos; outros duram horas, dias ou até meses. Essas variações podem produzir grandes diferenças na quantidade de proteína sintetizada.

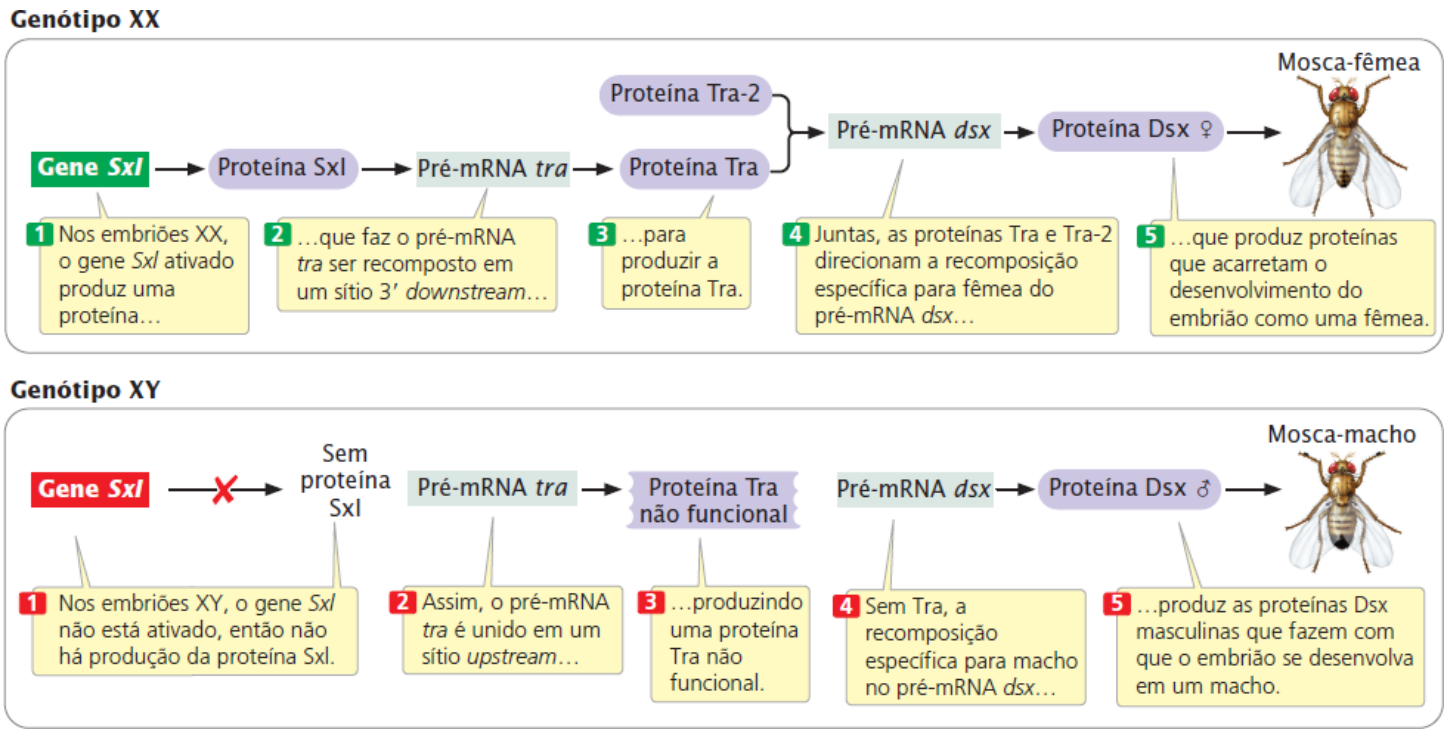


Figura 17.11 A recomposição alternativa controla a determinação do sexo na *Drosophila*.

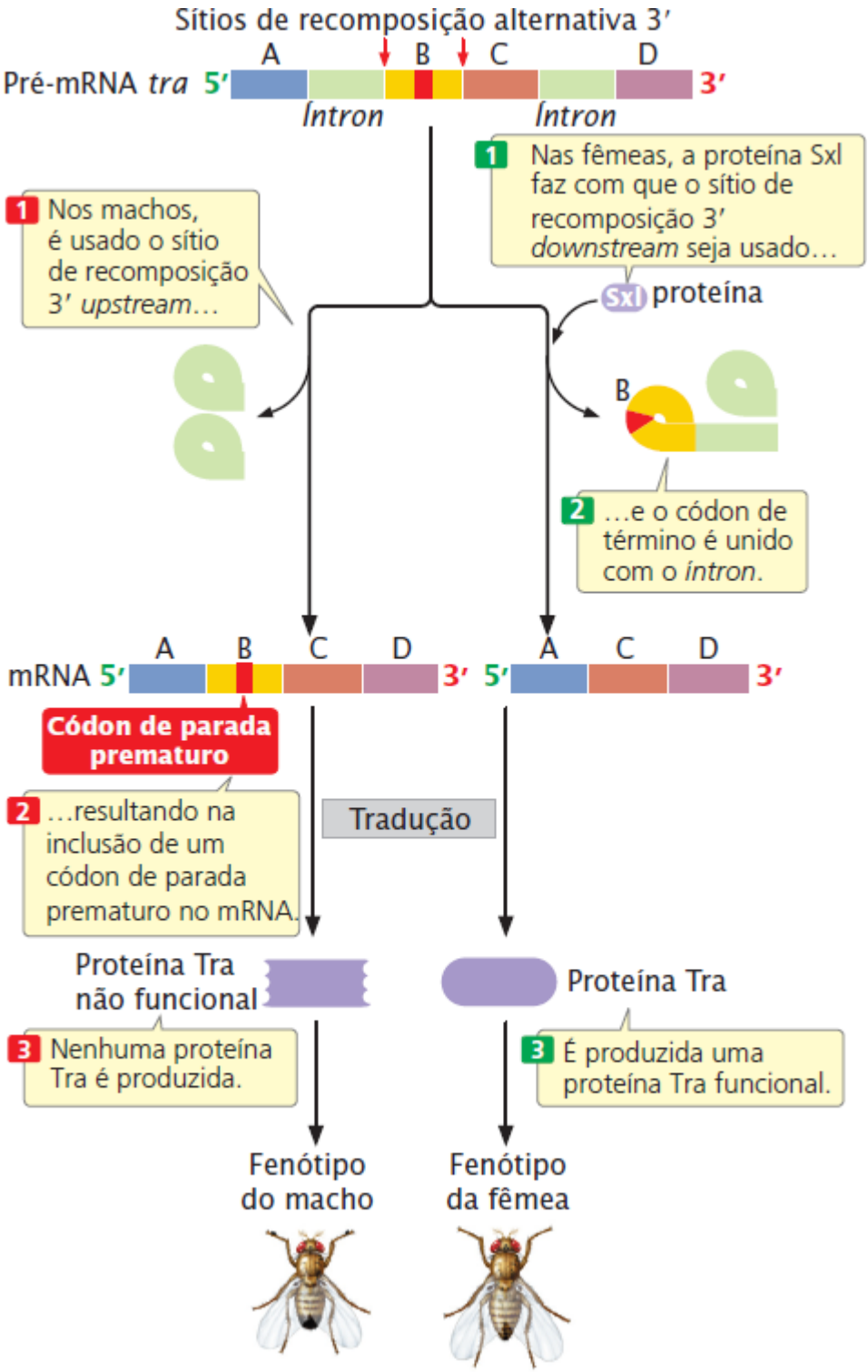


Figura 17.12 Recomposição alternativa do pré-mRNA *tra*. Existem dois sítios de recomposição alternativa 3′.

O RNA celular é degradado por ribonucleases, enzimas que degradam especificamente o RNA. A maioria das células eucarióticas tem 10 ou mais tipos de ribonucleases, e existem várias vias diferentes de degradação do mRNA. Em uma via, o cap 5′ é removido primeiro, seguido pela remoção 5′ → 3′ dos nucleotídios. Uma segunda via inicia-se na extremidade 3′ do mRNA e remove os nucleotídios no sentido 3′ → 5′. Em uma terceira via, o mRNA é rompido em sítios internos.

A degradação do RNA mensageiro a partir da extremidade 5′ é a mais comum e inicia com a remoção do cap 5′. Essa via é precedida pelo encurtamento da cauda poli(A). As proteínas ligadoras poli(A) se ligam normalmente à cauda poli(A) e contribuem para seu efeito acentuador da estabilidade. A existência dessas proteínas na extremidade 3′ do mRNA protege o cap 5′. Quando a cauda poli(A) é encurtada abaixo de um limite crítico, o cap 5′ é removido, e as nucleases degradam o mRNA ao remover os nucleotídios a partir da extremidade 5′. Essas observações sugerem que o cap 5′ e a cauda poli(A) do mRNA eucariótico interajam fisicamente um com outro – o mais provável é que a cauda poli(A) se dobre sobre si mesma de forma que os PABPs entrem em contato com o cap 5′ (ver Figura 15.16).

Parte da degradação do RNA ocorre em complexos especializados chamados corpúsculos P. Entretanto, parece que os corpúsculos P são mais do que apenas sítios de destruição para RNA. Evidências sugerem que os corpúsculos P podem

armazenar temporariamente as moléculas de mRNA, que mais tarde podem ser liberadas e traduzidas. Assim, os corpúsculos P ajudam a controlar a expressão dos genes ao regular quais moléculas de RNA são degradadas e quais são sequestradas para serem liberadas posteriormente. A degradação do RNA facilitada pelos pequenos RNAs interferentes (siRNAs) também ocorre nos corpúsculos P (ver próxima seção).

Outras partes do mRNA eucariótico, incluindo as sequências na região 5′ não traduzida (5′ UTR), a região codificadora e a 3′ UTR, também afetam a estabilidade do mRNA. Os mRNAs eucarióticos de vida curta têm uma ou mais cópias de uma sequência consenso composta por 5′-AUUUAUAA-3′, chamada de elemento rico em AU, na 3′ UTR. Os RNAs mensageiros contendo os elementos ricos em AU são degradados por um mecanismo no qual os microRNAs participam (ver próxima seção).

Resolva o Problema 26

Conceitos

A estabilidade do mRNA influencia a expressão gênica ao afetar a quantidade de mRNA disponível para ser traduzido. A estabilidade do mRNA é afetada pelo cap 5′, pela cauda poli(A) e pela seção codificadora 5′ UTR, e as sequências na 3′ UTR.

Checagem dos conceitos 4

Como a cauda poli(A) afeta a estabilidade?

17.5 A interferência por RNA é um mecanismo importante de regulação gênica

A expressão de vários genes eucarióticos é controlada por meio da interferência por RNA, também conhecida como silenciamento por RNA e silenciamento gênico pós-transcricional (ver *Capítulo 14*). A pesquisa sugere que cerca de 30% dos genes humanos são regulados por interferência por RNA. A interferência por RNA é comum nos eucariotos e ocorre em fungos, plantas e animais. Esse mecanismo também é muito empregado como técnica poderosa para regular artificialmente a expressão gênica em organismos geneticamente modificados (ver *Capítulo 19*).

RNAs de interferência pequenos e microRNAs

A interferência por RNA é acionada pelos microRNAs (miRNAs) e por RNAs de interferência pequenos (siRNAs), dependendo da origem e do modo de ação (ver *Capítulo 14*). Uma enzima chamada Dicer rompe e processa o RNA de fita dupla para produzir siRNAs de fita simples ou miRNAs que têm de 21 a 25 nucleotídios de comprimento (**Figura 17.13**) e pareia com proteínas para formar um complexo silenciador induzido pelo RNA (RISC). O componente RNA do RISC então pareia com as sequências de base complementares de moléculas de mRNA específicas, muito frequentemente com sequências na 3′ UTR do mRNA. RNAs de interferência pequenos tendem a parear perfeitamente com as bases dos mRNAs, enquanto os miRNAs formam pareamentos menos perfeitos.

Mecanismos de regulação gênica pela interferência por RNA

RNAs de interferência pequenos e microRNAs regulam a expressão gênica por meio de pelo menos quatro mecanismos distintos: (1) clivagem do mRNA, (2) inibição da tradução, (3) silenciamento da transcrição ou (4) degradação do mRNA.

Clivagem do RNA. Os RISCs que têm um siRNA (e alguns que têm um miRNA) pareiam com moléculas de mRNA e clivam o mRNA próximo ao meio do siRNA ligado (ver Figura 17.13 A). Essa clivagem é feita por uma proteína que às vezes é chamada de “Slicer”. Após a clivagem, o mRNA é degradado. Assim, quando há siRNAs e miRNAs, a velocidade de degradação dos mRNA aumenta e a quantidade de proteína produzida diminui.

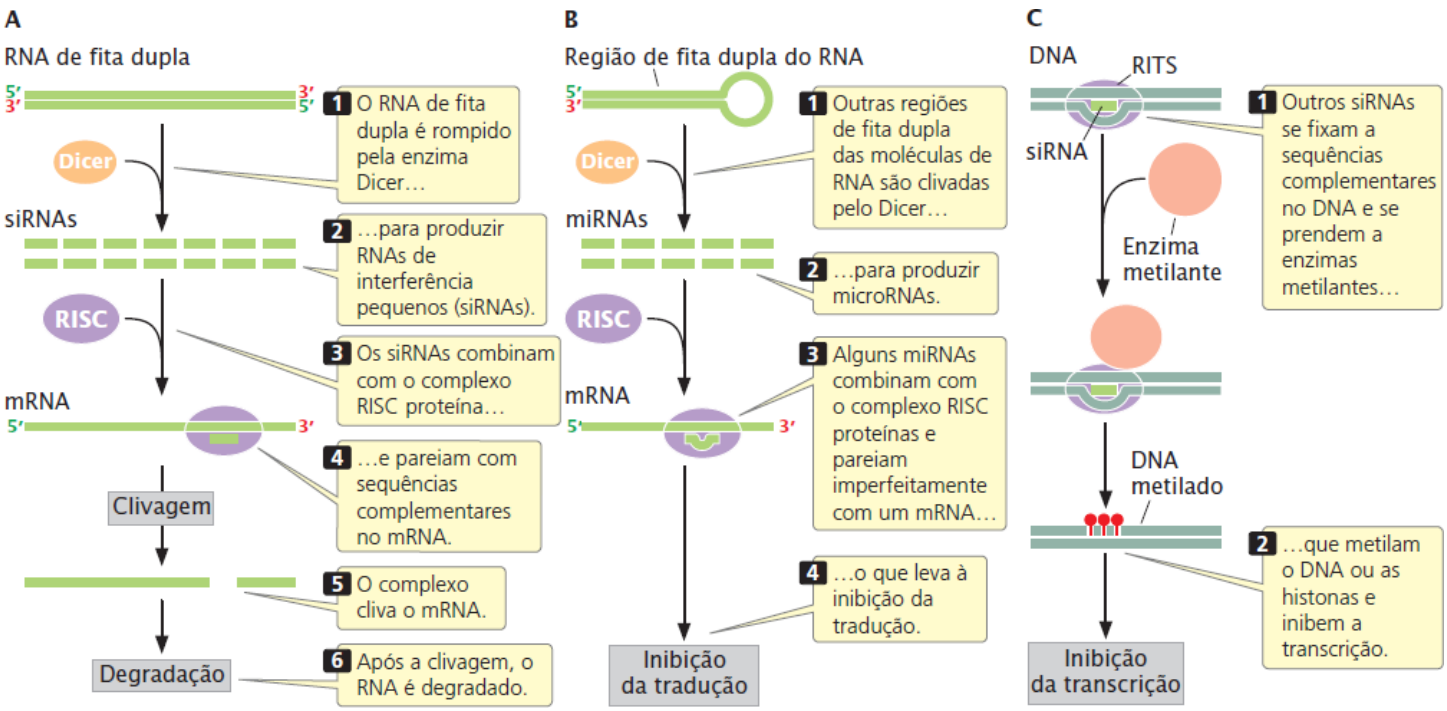


Figura 17.13 O silenciamento do RNA leva à degradação do mRNA ou à inibição da tradução ou transcrição. A. RNAs de interferência pequenos (*siRNAs*) degradam mRNA, rompendo-o. **B.** MicroRNAs (*miRNAs*) levam à inibição da tradução. **C.** Alguns RNAs de interferência pequenos (*siRNAs*) metilam as proteínas histona ou DNA, inibindo a transcrição.

Inibição da tradução. Alguns miRNAs regulam os genes que inibem a tradução dos seus mRNAs complementares (ver Figura 17.13 B). Por exemplo, um gene importante no desenvolvimento da flor na *Arabidopsis thaliana* é *APETALA2*. A expressão desse gene é regulada por um miRNA que pareia pares de base com nucleotídios na região codificadora do mRNA do *APETALA2* e inibe sua tradução.

O mecanismo exato pelo qual os miRNAs suprimem a tradução não é conhecido, mas a pesquisa sugere que ele iniba tanto a etapa de início da tradução e as etapas após o início da tradução, como as que provocam o término prematuro. Muitos mRNAs têm múltiplos sítios de ligação do miRNA, e a tradução é inibida mais eficientemente quando múltiplos miRNAs estão ligados ao mRNA.

Silenciamento transcricional. Outros siRNAs silenciam a transcrição ao alterar a estrutura da cromatina. Esses siRNAs se combinam com proteínas para formar um complexo chamado RITS (para silenciamento transcricional do RNA; ver Figura 17.13 C), que é análogo ao RISC. O componente siRNA do RITS se liga a essa sequência complementar no DNA ou uma molécula de RNA no processo de ser transcrito e reprime a transcrição ao atrair enzimas que metilam as caudas das proteínas histona. A adição dos grupos metila às histonas faz com que elas se liguem ao DNA mais firmemente, restringindo o acesso das proteínas e enzimas necessárias para fazer a transcrição (ver a seção anterior, sobre modificação da histona). Alguns miRNAs se ligam a sequências complementares no DNA e atraem enzimas que metilam o DNA diretamente, o que também leva à supressão da transcrição (ver a seção anterior, sobre metilação do DNA).

Degradação do mRNA independente de Slicer. Um mecanismo final pelo qual os miRNAs regulam a expressão gênica é a deflagração do decaimento do mRNA em um processo que não requer a atividade de Slicer. Por exemplo, um mRNA de vida curta com um elemento rico em AU na sua 3′ UTR é degradado por um mecanismo de silenciamento do RNA. Os pesquisadores identificaram um miRNA com uma sequência complementar à sequência consenso no elemento rico em AU. Esse miRNA se liga ao elemento rico em AU e, de um modo que ainda não é totalmente compreendido, leva à degradação do mRNA em um processo que requer Dicer e RISC.

Controle do desenvolvimento por interferência por RNA

Parte do desenvolvimento nos eucariotos multicelulares é feita por regulação gênica: diferentes genes são ligados e desligados em momentos específicos (ver *Capítulo 22*). De fato, quando os miRNAs foram descobertos, os pesquisadores perceberam que uma mutação em um miRNA em *C. elegans* provocava um defeito no desenvolvimento. Pesquisa recente demonstrou que as moléculas de miRNAs são fatores importantes no controle do desenvolvimento de plantas, animais e humanos. Por exemplo, o coração dos vertebrados se desenvolve por meio de diferenciação programada e proliferação dos cardiomiócitos, que são controlados por um miRNA específico chamado de miR-1-1.

Estudos recentes demonstraram que, por seus efeitos na expressão gênica, os miRNAs têm funções importantes em muitas

doenças e distúrbios. Por exemplo, uma forma genética de perda auditiva está associada a uma mutação no gene que codifica um miRNA. Outros miRNAs estão associados à cardiopatia. Um miRNA chamado miR-1-2 é expresso em grande quantidade no músculo cardíaco. Camundongos geneticamente modificados para expressar apenas 50% da quantidade normal de miR-1-2 frequentemente têm orifícios no septo interventricular, uma cardiopatia congênita comum em humanos recém-nascidos. A expressão excessiva de outro miRNA chamado miR-1 no coração de camundongos adultos provoca arritmia cardíaca – atividade elétrica irregular do coração, que é potencialmente fatal nos seres humanos. As mudanças na expressão de outros miRNAs estão associados ao câncer.

Conceitos

O silenciamento do RNA é iniciado por moléculas de RNA de fita dupla que eram clivadas e processadas. Os siRNAs ou miRNAs resultantes combinam com proteínas para formar complexos que se ligam a sequências complementares no mRNA ou no DNA. Os siRNAs e miRNAs afetam a expressão gênica ao clivar o mRNA, inibir a tradução, alterar a estrutura da cromatina ou acionar a degradação do RNA.

Checagem dos conceitos 5

No silenciamento do RNA, os siRNAs e miRNAs se ligam a qual parte das moléculas de mRNA que eles controlam?

a. 5′ UTR.
b. Segmento que codifica aminoácidos.
c. Cauda poli(A) 3′.
d. 3′ UTR.

17.6 Alguns genes são regulados por processos que afetam a tradução ou por modificações das proteínas

Os ribossomos, aminoacil tRNAs, fatores de iniciação e fatores de alongamento são necessários para a tradução das moléculas de mRNA. A disponibilidade desses componentes afeta a taxa de tradução e, portanto, influencia a expressão gênica. Por exemplo, a ativação dos linfócitos T (células T) é decisiva para o desenvolvimento das respostas imunológicas a vírus (ver *Capítulo 22*). Os linfócitos T estão normalmente no estágio G_0 do ciclo celular e não estão se dividindo ativamente. Ao serem expostos a antígenos virais, entretanto, linfócitos T específicos são ativados e sofrem rápida proliferação (**Figura 17.14**). A ativação inclui um aumento de 7 a 10 vezes na síntese de proteínas, que faz com que o linfócito T entre no ciclo celular e prolifere. Esse pulso de síntese de proteínas não demanda aumento na síntese de mRNA. Em vez disso, um aumento global na síntese de proteína ocorre por causa da maior disponibilidade dos fatores de iniciação que participam da tradução – os fatores de iniciação que possibilitam que os ribossomos se liguem ao mRNA e iniciem a tradução. Esse aumento nos fatores de iniciação leva a mais tradução a partir das moléculas de mRNA existentes, aumentando a quantidade geral de proteína sintetizada. Do mesmo modo, a insulina estimula o início da síntese de proteína ao aumentar a disponibilidade dos fatores de iniciação. Os fatores de iniciação existem nas formas inativas e, em resposta a vários sinais celulares, podem ser ativados por modificações químicas da sua estrutura, como a fosforilação.

Também existem mecanismos para a regulação da tradução de mRNAs específicos. O início da tradução em alguns mRNAs é regulado por proteínas que se ligam à 5′ UTR dos mRNAs e inibem a ligação dos ribossomos, semelhante à maneira na qual as proteínas repressoras se ligam aos operadores e impedem a transcrição dos genes estruturais. A tradução de alguns mRNAs é afetada pela ligação de proteínas a sequências na 3′ UTR.

Muitas proteínas eucarióticas são substancialmente modificadas após a tradução pela clivagem seletiva e pela retirada de aminoácidos das extremidades, pela acetilação ou pela adição de grupos fosfato, grupos carboxila, grupos metila ou carboidratos à proteína. Essas modificações afetam o transporte, a função e a atividade das proteínas.

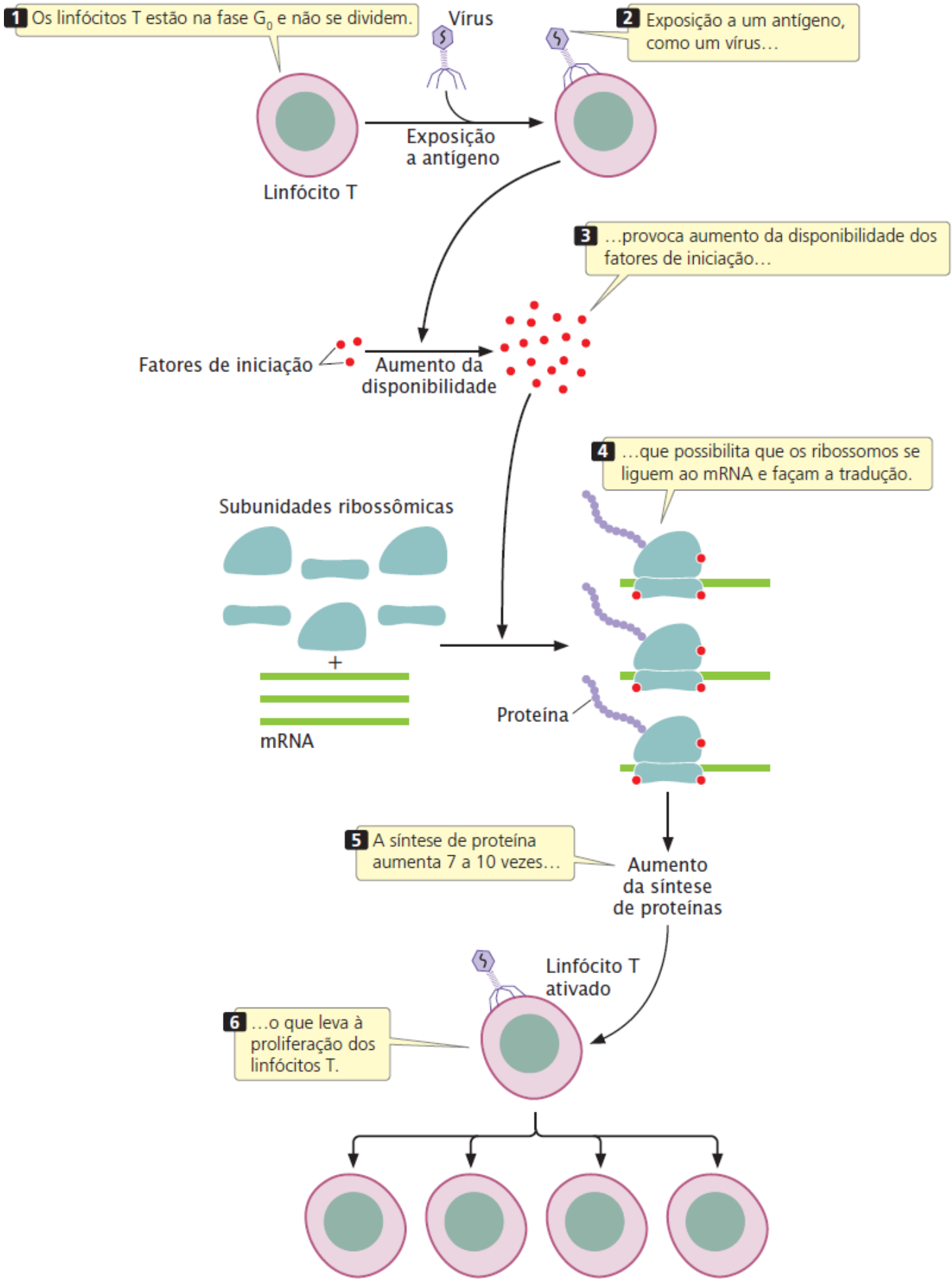


Figura 17.14 A expressão de alguns genes eucarióticos é regulada pela disponibilidade dos componentes necessários para a tradução. Neste exemplo, a exposição a um antígeno estimula uma disponibilidade maior de fatores de iniciação e um subsequente aumento da síntese de proteínas, levando à proliferação dos linfócitos T.

Conceitos

O início da tradução é afetado pelas proteínas que se ligam a sequências específicas na extremidade 5′ do mRNA. A disponibilidade de ribossomos, tRNAs, fatores de iniciação e de alongamento e outros componentes do aparato de tradução influencia a taxa de tradução. A tradução de alguns mRNAs é regulada por proteínas que se ligam a regiões 5′ e 3′ não traduzidas do mRNA.

Conceitos conectantes

Comparação do controle da expressão gênica bacteriana e eucariótica

Agora que analisamos os principais tipos de regulação gênica nas bactérias (ver no *Capítulo 16*) e nos eucariotos (ver neste capítulo), vamos analisar algumas das semelhanças e diferenças nos controles gênicos bacteriano e eucariótico.

1. Parte da regulação gênica nas células bacterianas está no nível da transcrição (embora essa regulação também ocorra em outros níveis). A regulação gênica nas células eucarióticas ocorre em múltiplos níveis, incluindo a estrutura da cromatina, transcrição, processamento do mRNA, estabilidade do RNA, interferência por RNA e controle pós-tradução.
2. Os complexos eventos bioquímicos e de desenvolvimento nas células bacterianas e eucarióticas precisam de uma cascata da regulação gênica, na qual a ativação de um conjunto de genes estimula a ativação de outro conjunto.
3. Parte da regulação gênica nas células bacterianas e eucarióticas é feita por meio de proteínas que se ligam a sequências específicas no DNA. Existem vários tipos de proteínas regulatórias, mas a maioria pode ser caracterizada de acordo com um pequeno conjunto de motivos de ligação ao DNA.
4. A estrutura da cromatina participa na regulação gênica eucariótica (mas não bacteriana). Em geral, a cromatina condensada reprime a expressão gênica; a estrutura da cromatina tem de ser alterada antes da transcrição. A estrutura da cromatina é alterada pelas proteínas remodeladoras de cromatina, pela modificação das proteínas histona e pela metilação do DNA.
5. As modificações na estrutura da cromatina nos eucariotos podem levar a mudanças epigenéticas, que afetam a expressão gênica e são transmitidas para outras células ou para futuras gerações.
6. Nas células bacterianas, os genes estão frequentemente agrupados em óperons e são expressos de forma coordenada por meio da transcrição em uma única molécula de mRNA. Em contrapartida, muitos genes eucarióticos têm seus próprios promotores e são transcritos de forma independente. A regulação coordenada nas células eucarióticas ocorre frequentemente graças a elementos de resposta comuns, encontrados nos promotores e acentuadores dos genes. Diferentes genes que têm o mesmo elemento de resposta são influenciados pela mesma proteína regulatória.
7. As proteínas regulatórias que afetam a transcrição exibem dois tipos básicos de controle: os *repressores* inibem a transcrição (controle negativo); os *ativadores* estimulam a transcrição (controle positivo). As células bacterianas e eucarióticas têm controles positivo e negativo.
8. O início da transcrição é um processo relativamente simples nas células bacterianas, e as proteínas regulatórias bloqueiam ou estimulam a ligação da RNA polimerase ao DNA. Por outro lado, a transcrição eucariótica demanda um maquinário complexo que inclui RNA polimerase, fatores de transcrição gerais e ativadores da transcrição, que fazem com que a transcrição seja influenciada por múltiplos fatores.
9. Algumas proteínas regulatórias da transcrição dos eucariotos atuam a uma distância do gene ao se ligar aos acentuadores, levando à formação de uma alça no DNA que aproxima o promotor e o acentuador. Já foram descritas algumas sequências de ação a distância análogas aos acentuadores nas células bacterianas, mas parecem ser menos comuns.
10. O intervalo de tempo maior entre a transcrição e a tradução nas células eucarióticas comparado com o intervalo nas células bacterianas permite que a estabilidade do mRNA e o processamento do mRNA tenham participações maiores na regulação gênica eucariótica.
11. As moléculas de RNA (RNA antissenso) atuam como reguladores da expressão gênica nas bactérias. As células bacterianas não apresentam regulação por siRNAs e miRNAs, que é significativa nos eucariotos.

O **Quadro 17.2** resume as semelhanças e diferenças na regulação gênica nas bactérias e nos eucariotos.

Quadro 17.2 Comparação da expressão gênica nas bactérias e nos eucariotos.

Característica	Controle gênico bacteriano	Controle gênico eucariótico
Níveis de regulação	Principalmente transcrição	Muitos níveis
Cascatas de regulação gênica	Presentes	Presentes
Proteínas ligadoras do DNA	Importantes	Importantes
Papel da estrutura da cromatina	Ausente	Importante
Existência de óperons	Comum	Incomum
Controle negativo e positivo	Presente	Presente
Início da transcrição	Relativamente simples	Relativamente complexo
Acentuadores	Menos comuns	Mais comuns
Transcrição e tradução	Ocorrem simultaneamente	Ocorrem separadamente
Regulação por pequenos RNAs	Rara	Comum

Resumo dos conceitos

- As células eucarióticas diferem das bacterianas em várias maneiras que afetam a regulação gênica, incluindo, nos eucariotos, a ausência de óperons e a ocorrência de cromatina e de uma membrana nuclear
- Nas células eucarióticas, a estrutura de cromatina reprime a expressão gênica. Na transcrição, a estrutura da cromatina é alterada pelo reposicionamento dos nucleossomos e pela modificação das proteínas histona, incluindo acetilação, fosforilação e metilação. A metilação do DNA também afeta a transcrição
- O início da transcrição eucariótica é controlado por fatores de transcrição gerais que se depositam no aparato basal de transcrição e por proteínas reguladoras de transcrição que estimulam ou reprimem os níveis normais de transcrição ao se ligar a promotores regulatórios e acentuadores
- Acentuadores afetam a transcrição de genes distantes. Proteínas regulatórias se ligam a acentuadores e interagem com o aparato basal de transcrição ao fazer o DNA interveniente se soltar
- Sequências de DNA chamadas insuladores limitam a ação dos acentuadores ao bloquear sua ação de modo dependente da posição
- Alguns fatores regulatórios frazem a RNA polimerase pare *downstream* do promotor
- Os genes controlados de forma coordenada nas células eucarióticas respondem aos mesmos fatores porque eles têm elementos de resposta em comum que são estimulados pelo mesmo ativador da transcrição
- A expressão gênica nas células eucarióticas pode ser influenciada pelo processamento do RNA
- A expressão gênica pode ser regulada por mudanças na estabilidade do RNA. O cap 5′, a sequência codificadora, a 3′ UTR e a cauda poli(A) são importantes no controle da estabilidade dos mRNAs eucarióticos. As proteínas que se ligam às extremidades 5′ e 3′ do mRNA eucariótico podem afetar sua tradução
- O silenciamento do RNA é importante na regulação gênica eucariótica. Pequenas moléculas de RNA (siRNAs e miRNAs) clivadas do RNA de fita dupla se combinam com proteínas e se ligam a sequências de mRNA ou DNA. Esses complexos rompem o RNA, inibem a tradução, afetam a degradação do RNA e silenciam a transcrição
- O controle da modificação pós-tradução das proteínas pode participar da expressão gênica.

Termos importantes

Código da histona
Complexo de remodelagem da cromatina
Elemento de resposta
Ilha CpG
Insulador
Mediador

Proteínas de choque térmico
Proteínas SR
Sítios hipersensíveis da DNAse I

Respostas da Checagem dos conceitos

1. Os três processos gerais são remodelagem da cromatina, modificação das proteínas histona (p. ex., metilação e acetilação das histonas) e metilação do DNA.
2. b.
3. O DNA entre o acentuador e promotor se solta, de modo que as proteínas regulatórias ligadas ao acentuador conseguem interagir diretamente com o aparato de transcrição.
4. A cauda poli(A) estabiliza o cap 5′, que tem de ser removido antes que a molécula de mRNA possa ser degradada a partir da extremidade 5′.
5. d.

Problema desenvolvido

Problema

Qual seria o efeito de uma mutação que tornou as proteínas ligadoras da cauda poli(A) não funcionais?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Uma afirmação do efeito de uma mutação que eliminou a função da proteína ligadora da poli(A).

Quais informações são fornecidas para solucionar o problema?
- Ocorre uma mutação no gene que codifica a proteína ligadora de poli(A)
- A mutação faz com que a proteína poli(A) seja não funcional.

Para a solução deste problema, revise:
Degradação do RNA, na Seção 17.4.
Adição da cauda poli(A), na Seção 14.2 (*Capítulo 14*).

Etapas para a solução

O RNA mensageiro pode ser degradado a partir da extremidade 5′, da extremidade 3′ ou por clivagem interna. A degradação a partir da extremidade 5′ exige a remoção do cap 5′ e, em geral, é precedida pelo encurtamento da cauda poli(A). As proteínas ligadoras da poli(A) se conectam normalmente à cauda poli(A) e evitam que ela seja encurtada. Assim, essas proteínas na cauda poli(A) protegem o cap 5′, o que evita a degradação do RNA. Se o gene para as proteínas ligadoras da poli(A) sofrer mutação de modo que sejam produzidas proteínas poli(A) não funcionais, as proteínas não se ligariam à cauda poli(A). A cauda seria encurtada prematuramente, o cap 5′ seria removido e o mRNA seria degradado com mais facilidade. O resultado final seria menos mRNA e muito menos síntese de proteína.

Lembrete: A cauda poli(A) afeta a estabilidade do mRNA.

Questões de compreensão

Introdução

1. Quais são as semelhanças entre os humanos e os chimpanzés? Quais mudanças genéticas poderiam ser responsáveis pelas grandes diferenças na anatomia, fisiologia e comportamento dos humanos e chimpanzés?

Seção 17.1

2. Liste algumas diferenças importantes entre células bacterianas e eucarióticas que afetam a regulação gênica.

Seção 17.2

3. Onde os sítios hipersensíveis da DNase I são encontrados e o que eles indicam sobre a natureza da cromatina?
4. Que mudanças ocorrem na estrutura da cromatina e qual é a participação dessas mudanças na regulação gênica eucariótica?
5. O que é o código da histona?
6. Como a imunoprecipitação da cromatina é usada para determinar as localizações das modificações da histona no genoma?

Seção 17.3

7. Explique resumidamente como as proteínas ativadoras da transcrição e repressores afetam o nível de transcrição dos genes eucarióticos.
8. O que é um acentuador? Como ele afeta a transcrição de genes distantes?
9. O que é um insulador?
10. O que é um elemento de resposta? Como os elementos de resposta fazem a expressão coordenada dos genes eucarióticos?

Seção 17.4

11. Destaque o papel da recomposição (*splicing*) alternativa no controle da diferenciação do sexo na *Drosophila*.
12. Qual é o papel da estabilidade do RNA na regulação gênica? O que controla a estabilidade do RNA nas células eucarióticas?

Seção 17.5

13. Liste resumidamente algumas das maneiras pelas quais os siRNAs e os miRNAs regulam os genes.

Seção 17.6

14. Como a regulação gênica bacteriana difere da regulação gênica eucariótica? E quais são as semelhanças?

Questões e problemas aplicados

Seção 17.2

15. A malária, uma das doenças infecciosas mais disseminadas e destruidoras, é causada pelos parasitas protozoários do gênero *Plasmodium*, que são transmitidos entre as pessoas por mosquitos. O parasita *Plasmodium* se esquiva do sistema imunológico do hospedeiro ao alternar constantemente a expressão de seus genes *var*, que codificam os antígenos de superfície do *Plasmodium* (L. H. Freitas-Junior *et al*. 2005. *Cell* 121:25-36). Os genes *var* são expressos quando a estrutura da cromatina é fragmentada por mudanças químicas nas proteínas histona. Que tipo de mudanças químicas nas proteínas histona poderia ser responsável por essas mudanças na expressão gênica?
16. Um geneticista está tentando determinar quantos genes são encontrados em uma região de 300.000 pb do DNA. A análise mostra que foram encontradas quatro modificações H3K4me3 nesse segmento de DNA. O que isso poderia sugerir sobre o número de genes localizados nesse segmento?
17. Em uma linhagem de células humanas obtidas em cultura, um geneticista isola uma mutação termossensível em um *locus* que codifica uma enzima acetiltransferase; em temperaturas acima de 38°C, as células mutantes produzem uma forma não funcional da enzima. Qual seria o efeito mais provável dessa mutação quando as células crescem a 40°C?

*18. Qual seria o efeito mais provável da deleção do *locus D da floração* (*FLD*) na *Arabidopsis thaliana*?

*19. X31b é um composto experimental captado pelas células em rápida divisão. Pesquisa mostrou que X31b estimula a metilação do DNA. Alguns pesquisadores de câncer estão interessados em testar X31b para tratar o câncer de próstata. Ofereça uma possível explicação para o X31b ser um agente antineoplásico efetivo.

Seção 17.3

20. Qual é a diferença entre os repressores que se ligam aos silenciadores e os eucariotos dos repressores que se ligam aos operadores nas bactérias?
21. Examine a **Figura 17.7**. Qual seria o efeito na transcrição se ocorresse uma mutação no gene que codifica GAL3, de modo que não fosse produzida GAL3 funcional?
22. Qual seria o efeito de deslocar o insulador apresentado na **Figura 17.8** para uma posição entre o Acentuador II e o promotor para o gene B?

*23. Um acentuador é circundado por quatro genes (*A*, *B*, *C* e *D*), como apresentado no diagrama a seguir. Existe um insulador

entre o gene *C* e o gene *D*. Com base nas posições dos genes, do acentuador e do insulador, a transcrição de quais genes tem maior probabilidade de ser estimulada pelo acentuador? Explique sua justificativa.

Seção 17.4

*24. Qual será o efeito no desenvolvimento sexual em embriões recém-fertilizados de *Drosophila* se os seguintes genes forem eliminados?
   a. *Sexo letal*.
   b. *Transformador*.
   c. *Sexo duplo*.

25. Examine a **Figura 17.12**. Qual seria o efeito de uma mutação que eliminasse o segundo sítio de corte 3′ na extremidade do éxon B no pré-mRNA *tra*?

*26. Alguns mRNAs eucarióticos têm um elemento rico em AU na região 3′ não traduzida. Qual seria o efeito sobre a expressão gênica se esse elemento sofresse mutação ou deleção?

27. Uma cepa de *Arabidopsis thaliana* tem uma mutação no gene *APETALA2*, no qual parte da região 3′ não traduzida do mRNA transcrito a partir do gene é eliminada. Qual seria o efeito mais provável dessa mutação na expressão do gene *APETALA2*?

28. Qual será o efeito de uma mutação que destrói a capacidade da proteína ligadora de poli(A) para se fixar à cauda poli(A)?

Seção 17.5

29. Suponha que um geneticista introduza um pequeno RNA interferente (siRNA) complementar ao mRNA FLC na **Figura 17.3**. Qual seria o efeito na floração da *Arabidopsis*? Explique sua resposta.

Questões desafiadoras

Seção 17.3

ANÁLISE DE DADOS

30. O gene *SER3* da levedura, que participa da biossíntese da serina, é reprimido durante o crescimento em meio de cultura rico em nutriente, de modo que ocorre pouca transcrição e pouca enzima SER3 é produzida. Em um estudo da repressão do gene *SER3*, foi descoberto que uma região do DNA *upstream* do *SER3* era intensamente transcrita quando o gene *SER3* era reprimido (J. A. Martens, L. Laprade, and F. Winston. 2004. *Nature* 429:571-574). Nesta região *upstream* existe um promotor que estimula a transcrição de uma molécula de RNA chamada *SRG1* RNA (para o gene regulatório 1 *SER3*). Essa molécula de RNA não tem as sequências necessárias para a tradução. As mutações no promotor para *SRG1* levam ao desaparecimento do *SRG1* RNA, e essas mutações removem a repressão do *SER3*. Quando a RNA polimerase se liga ao promotor *SRG1*, a polimerase se desloca *downstream*, transcrevendo o RNA *SGR1* e passando e transcrevendo o promotor para *SER3*. Esta atividade leva à repressão do *SER3*. Proponha uma possível explicação para como a transcrição do *SGR1* poderia reprimir a transcrição do *SER3*. (Dica: Lembre-se de que o RNA *SGR1* não codifica uma proteína.)

Seção 17.5

31. Uma característica comum de muitos mRNAs eucarióticos é uma 3′UTR algo longa que tem sequências consenso. A creatinoquinase B (CK-B) é uma enzima importante no metabolismo celular. Algumas células – chamadas de células U937D – têm muito mRNA de CK-B, mas não há enzima CK-B. Nessas células, a extremidade 5′ do mRNA de CK-B está ligada a ribossomos, mas o mRNA aparentemente não é traduzido. Algo inibe a tradução do mRNA de CK-B nessas células.

    Os pesquisadores introduziram numerosos segmentos curtos de RNA contendo apenas sequências 3′UTR nas células U937D. Como resultado, as células U937D começaram a sintetizar a enzima CK-B, mas a quantidade total de mRNA de CK-B não aumentou. A introdução de segmentos curtos de outras sequências de RNA não estimulou a síntese de CK-B; apenas as sequências 3′UTR acionaram a tradução da enzima.

    Com base nesses resultados, proponha um mecanismo para explicar como a tradução de CK-B é inibida nas células U937D. Explique como a introdução de segmentos curtos de RNA com sequências 3′UTR poderia anular a inibição.

18 Mutações do Gene e Reparo do DNA

Uma mosca sem coração

O coração de uma mosca-da-fruta é um órgão simples, um tubo aberto que se contrai de forma rítmica, bombeando o fluido – ação algo ineficiente – para o corpo da mosca. Embora simples e deselegante, o coração da mosca-da-fruta é essencial. Algumas moscas mutantes raras nunca desenvolvem um coração e morrem (o que seria esperado) em um estágio embrionário inicial. O geneticista Rolf Bodmer analisou esses mutantes em 1980 e fez uma importante descoberta – um gene que especifica o desenvolvimento de um coração. Ele chamou o gene de *tinman* (homem de lata), em homenagem ao personagem do Mágico de Oz que também não tinha coração. A pesquisa de Bodmer revelou que o *tinman* codifica um fator de transcrição que se liga ao DNA e liga outros genes essenciais para o desenvolvimento normal de um coração. Nas moscas mutantes sem esse gene, o fator de transcrição nunca era produzido e o coração não se desenvolvia. Os achados de pesquisa subsequente revelaram a existência de um gene humano (chamado *Nkx2*.5) com uma sequência semelhante à do *tinman*, mas a função do gene humano é desconhecida.

Então, nos anos 1990, os médicos Jonathan e Christine Seidman começaram a estudar pessoas que nasciam com corações anormais, como os casos de coração com um orifício no septo que separa as câmaras dos lados direito e esquerdo do coração. Tais defeitos provocam fluxo sanguíneo anormal pelo coração, fazendo com que o músculo trabalhe mais que o normal e misturando o sangue oxigenado com o sangue não oxigenado. As cardiopatias congênitas não são incomuns; são encontradas em cerca de 1 em cada 125 recém-nascidos. Alguns defeitos desaparecem por si sós, mas outros precisam de cirurgia corretiva. Embora a cirurgia seja um sucesso para reverter esses problemas congênitos, muitos desses pacientes começam a apresentar arritmia cardíaca aos 20 ou 30 anos. Os Seidman e seus colaboradores descobriram várias famílias nas quais as cardiopatias congênitas e arritmias eram herdadas juntas em um padrão autossômico dominante. A análise molecular detalhada de uma dessas famílias revelou que o gene responsável pelas cardiopatias estava localizado no cromossomo 5, em um ponto onde o gene *tinman* humano (*Nkx2*.5) havia sido previamente mapeado. Todos os membros daquela família que herdavam as cardiopatias também herdavam uma mutação no gene *tinman*. Estudos adicionais com outros pacientes descobriram que muitas pessoas com cardiopatias congênitas tinham uma mutação nesse gene. A versão humana desse gene, assim como sua contrapartida nas moscas, codifica um fator de transcrição que controla o desenvolvimento do coração. Apesar das incríveis diferenças no tamanho, na anatomia e na fisiologia, os seres humanos e as moscas usam o mesmo gene para criar um coração.

A história do *tinman* ilustra a importância central do estudo das mutações: o exame dos mutantes é, em muitos casos, uma fonte de discernimento crítico nos processos biológicos importantes. Este capítulo é dedicado às mutações gênicas – como os erros surgem nas instruções genéticas e como esses erros são estudados. Vamos começar com um rápido exame dos diferentes tipos de mutações, incluindo seus efeitos fenotípicos, como eles podem ser suprimidos e as taxas de mutação. A seção a seguir explora como as mutações surgem espontaneamente durante e após a replicação do DNA, assim como as substâncias químicas e a radiação induzem as mutações. Após discutir a análise das mutações, examinaremos os elementos de transposição, as sequências de DNA capazes de se deslocar no genoma e, assim, produzir mutações. Finalmente, estudaremos o reparo do DNA e algumas das doenças que surgem quando ele é defeituoso.

18.1 As mutações são alterações herdadas na sequência do DNA

O DNA é uma molécula muito estável, replicada com incrível acurácia (ver os *Capítulos 10* e *12*), mas podem ocorrer mudanças na estrutura do DNA e erros de replicação. Uma **mutação** é definida como uma mudança herdada na informação genética; os descendentes podem ser células ou organismos.

Importância das mutações

As mutações são responsáveis tanto por manter a vida quanto por causar grande sofrimento. Por um lado, a mutação é a fonte de todas as variações genéticas, a matéria-prima da evolução. A capacidade dos organismos em se adaptar à mudança ambiental depende da variação genética nas populações naturais, e a variação genética é produzida pela mutação. Por outro lado, muitas mutações têm efeitos prejudiciais e são a causa de muitas doenças e distúrbios.

Parte do estudo da genética se concentra em como as variantes produzidas pela mutação são herdadas; os cruzamentos genéticos são insignificantes se todos os membros de uma espécie são identicamente homozigotos para os mesmos alelos. Parte do sucesso de Gregor Mendel em solucionar os princípios da hereditariedade pode ser atribuído ao uso de variantes cuidadosamente selecionadas de ervilha de jardim. Da mesma forma, Thomas Hunt Morgan e seus estudantes descobriram muitos princípios básicos de genética ao analisar moscas-da-fruta mutantes.

As mutações também são úteis para examinar processos biológicos fundamentais. Encontrar ou criar mutações que afetam diferentes componentes de um sistema biológico e estudar seus efeitos pode levar à compreensão do sistema. Esse método, chamado de dissecção genética, é como compreender como um automóvel funciona ao quebrar diferentes partes de um carro e observar os efeitos; por exemplo, ao esmagar o radiador, a máquina superaquece, revelando que o radiador resfria a máquina. O uso de mutações para comprometer o funcionamento pode, da mesma maneira, ser uma fonte de discernimento nos processos biológicos. Por exemplo, os geneticistas começaram a solucionar detalhes moleculares do desenvolvimento ao estudar as mutações, como o *tinman*, que interrompe vários estágios embrionários na *Drosophila* (ver *Capítulo 22*). Os cientistas também usaram análise de mutações para revelar as diferentes partes do óperon *lac* (discutido no *Capítulo 16*) e como eles atuam na regulação gênica. Embora quebrar "partes" para determinar sua função possa parecer uma abordagem grosseira para compreender um sistema, na verdade é uma ferramenta muito poderosa e tem sido usada em bioquímica, biologia do desenvolvimento, fisiologia e ciência comportamental. No entanto, esse método não é recomendado para aprender como seu carro funciona!

Conceitos

As mutações são mudanças herdadas no DNA. Elas são essenciais para o estudo da genética e são úteis em muitos outros campos biológicos.

Categorias de mutações

Nos organismos multicelulares, podemos diferenciar entre duas grandes categorias de mutações: mutações somáticas e mutações germinativas. As **mutações somáticas** surgem nos tecidos somáticos, que não produzem gametas (**Figura 18.1**). Quando uma célula somática com uma mutação se divide (mitose) a mutação é transmitida para suas células-filhas, gerando uma população de células geneticamente idênticas (um clone). Quanto mais precocemente uma mutação somática ocorrer, maior o clone de células que terão a mutação.

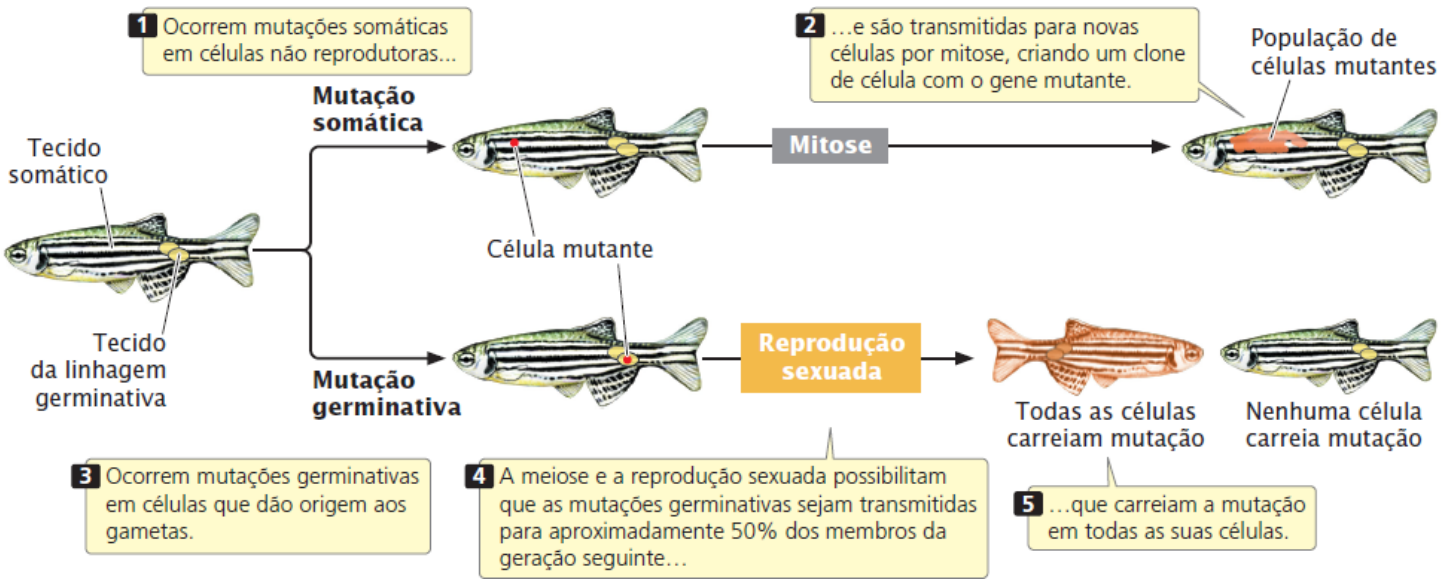


Figura 18.1 As duas classes básicas de mutações são mutações somáticas e mutações germinativas.

Em razão do grande número de células em um típico organismo eucariótico, as mutações somáticas são numerosas. Por

exemplo, existem cerca de 10^{14} células no corpo humano. Uma mutação surge uma vez a cada um milhão de divisões celulares, então centenas de milhões de mutações somáticas surgem em cada pessoa. Muitas mutações somáticas não têm efeito claro no fenótipo do organismo, porque a função da célula mutante é substituída pela das células normais ou a célula mutante morre e é substituída por células normais. Entretanto, as células com uma mutação somática que estimula a divisão celular podem aumentar o número e se disseminar; esse tipo de mutação pode dar origem a células com uma vantagem seletiva e é a base para os cânceres (ver *Capítulo 23*).

Mutações germinativas surgem em células que dão origem aos gametas. Uma mutação germinativa pode ser transmitida para futuras gerações, produzindo organismos que carreiam a mutação em todas as suas células somáticas e da linhagem germinativa (ver Figura 18.1). Quando falamos de mutações em organismos multicelulares, em geral estamos falando de mutações germinativas.

Historicamente, as mutações foram divididas entre as que afetam um único gene, chamadas de *mutações gênicas*, e as que afetam o número ou a estrutura dos cromossomos, chamadas de *mutações cromossômicas*. Essa distinção surgiu porque as mutações cromossômicas poderiam ser observadas diretamente ao microscópio, enquanto as mutações gênicas só poderiam ser detectadas pela observação de seus efeitos fenotípicos. Atualmente, o sequenciamento de DNA torna possível a observação direta das mutações gênicas, e as mutações cromossômicas são diferenciadas dessas arbitrariamente com base no tamanho da lesão do DNA. Contudo, é prático usar *mutação cromossômica* para descrever a alteração genética de grande escala que afeta a estrutura do cromossomo ou o número de cromossomos e usar **mutação gênica** para uma lesão de DNA relativamente pequena que afeta um único gene. Este capítulo discute as mutações gênicas; as mutações cromossômicas foram discutidas no Capítulo 8.

Tipos de mutações gênicas

Existem várias formas para classificar as mutações gênicas. Alguns esquemas de classificação baseiam-se na natureza do efeito fenotípico, outros se baseiam no agente causador da mutação e outros ainda se voltam para a natureza molecular do defeito. Aqui, vamos classificar as mutações principalmente com base na sua natureza molecular, mas também vamos encontrar alguns termos que relacionam as causas e os efeitos fenotípicos das mutações.

Substituições de base. O tipo mais simples de mutação gênica é uma **substituição de base**, a alteração de um único nucleotídio no DNA (**Figura 18.2 A**). Existem dois tipos de substituições de base. Em uma **transição**, uma purina é substituída por uma purina diferente ou uma pirimidina é substituída por uma pirimidina diferente (**Figura 18.3**). Em uma **transversão**, uma purina é substituída por uma pirimidina ou uma pirimidina é substituída por uma purina. O número de possíveis transversões (ver Figura 18.3) é duas vezes o número de possíveis transições, mas as transições surgem com maior frequência porque transformar uma purina em uma purina diferente ou uma pirimidina em uma pirimidina diferente é mais fácil do que transformar uma purina em uma pirimidina ou vice-versa. **Resolva o Problema 18**

Inserções e deleções. Outra classe de mutações gênicas inclui as **inserções** e **deleções** (coletivamente chamadas *indels*) – a adição ou a remoção, respectivamente, de um ou mais pares de nucleotídios (**Figura 18.2 B** e **C**). Embora as substituições de base sejam consideradas o tipo mais comum de mutação, a análise molecular revelou que as inserções e deleções são mais frequentes. As inserções e deleções nas sequências codificadoras de proteínas podem levar a **mutações no quadro de leitura (*frameshift*)**, mudanças de troca de quadro (ver *Capítulo 15*) do gene. As mutações no quadro de leitura em geral alteram todos os aminoácidos codificados pelos nucleotídios após a mutação e, então, elas têm efeitos drásticos no fenótipo. Algumas mutações de troca de quadro também introduzem códons de parada prematuros, interrompendo prematuramente a síntese de proteína e resultando em proteína encurtada (truncada). Nem todas as inserções e deleções levam a mutações de troca de quadro, entretanto; as inserções e deleções compostas por múltiplos de três nucleotídios deixam o quadro de leitura intacto, embora a adição ou remoção de um ou mais aminoácidos ainda possa afetar o fenótipo. *Indels* que não afetam o quadro de leitura são chamados de **inserções no quadro de leitura** e **deleções no quadro de leitura**.

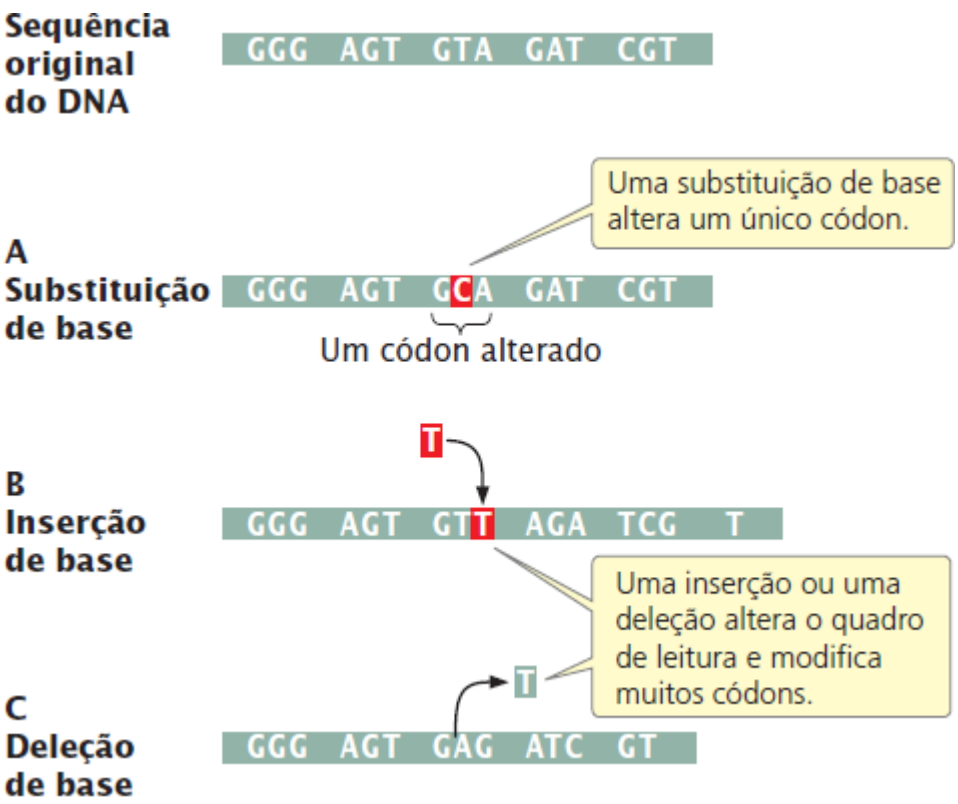


Figura 18.2 **Os três tipos básicos de mutações gênicas são substituições, inserções e deleções de base.**

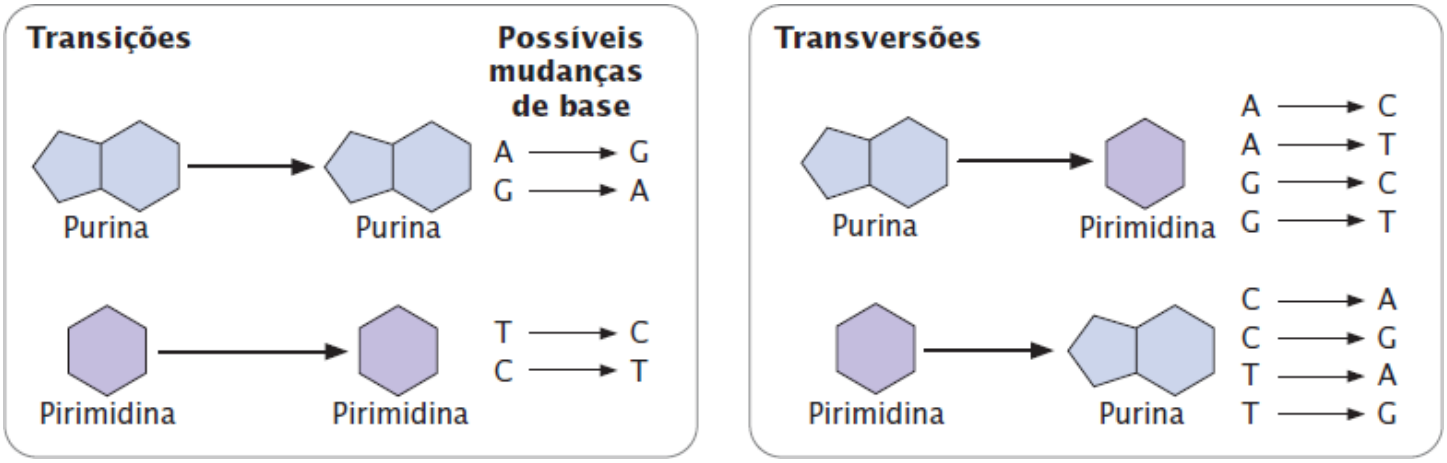


Figura 18.3 **Uma transição é a substituição de uma purina por uma purina ou de uma pirimidina por uma pirimidina; uma transversão é a substituição de uma pirimidina por uma purina ou de uma purina por uma pirimidina.**

Conceitos

As mutações gênicas são compostas por mudanças em um único gene e podem ser substituições de base (um único par de nucleotídios é alterado) ou inserções ou deleções (os nucleotídios são adicionados ou removidos). Uma substituição de base pode ser uma transição (substituição de bases semelhantes) ou uma transversão (substituição de bases diferentes). As inserções e deleções levam a alteração no quadro de leitura de um gene.

Checagem dos conceitos 1

Qual das mudanças a seguir é uma substituição de base por transição?

a. Adenina é substituída por timina.
b. Citosina é substituída por adenina.
c. Guanina é substituída por adenina.
d. Três pares de nucleotídios são inseridos no DNA.

Repetições expandidas de nucleotídios. As mutações nas quais o número de cópias de um conjunto de nucleotídios

aumenta são chamadas de **repetições expandidas de nucleotídios**. Esse tipo de mutação foi observado pela primeira vez em 1991 em um gene chamado *FMR-1*, que provoca a síndrome do X frágil, a causa hereditária mais comum de déficit intelectual. O distúrbio tem esse nome porque, em células especialmente tratadas de pessoas com a condição, a ponta de cada braço longo do cromossomo X é ligada a uma parte aparentemente delgada do cromossomo. O alelo *FMR-1* normal (sem a mutação) tem 60 ou menos cópias de CGG, mas em pessoas com a síndrome do X frágil, o alelo abriga centenas ou até milhares de cópias.

Já foram encontradas repetições expandidas de nucleotídios em quase 30 doenças humanas, muitas das quais estão listadas no **Quadro 18.1**. A maioria dessas doenças é causada pela expansão de um conjunto de três nucleotídios (chamado de trinucleotídio), em sua maioria CNG, onde N pode ser qualquer nucleotídio. Entretanto, algumas doenças são causadas por repetições de quatro, cinco e até 12 nucleotídios. O número de cópias de repetição de nucleotídios se correlaciona frequentemente com a gravidade ou a idade de início da doença. O número de cópias da repetição também tem relação com a instabilidade das repetições de nucleotídios: quando ocorrem mais repetições, a probabilidade de expansão para mais repetições aumenta. Essa associação entre o número de cópias de repetições de nucleotídios, a gravidade da doença e a probabilidade de expansão leva a um fenômeno conhecido como antecipação (ver *Capítulo 5*), no qual as doenças causadas por expansões repetidas de nucleotídios se tornam mais graves em cada geração. Mais raramente, o número de repetições de nucleotídios diminui em uma família. Também já foram observadas repetições expandidas de nucleotídios em alguns micróbios e plantas.

Aumentos do número de repetições de nucleotídios podem provocar sinais/sintomas da doença de várias maneiras. Em várias das doenças (p. ex., doença de Huntington), o nucleotídio se expande na parte codificadora de um gene, produzindo uma proteína tóxica que tem resíduos extras de glutamina (o aminoácido codificado pelo CAG). Em outras doenças, a repetição está fora da região codificadora de um gene e afeta sua expressão. Na síndrome do X frágil, cópias adicionais da repetição de nucleotídio fazem com o que o DNA seja metilado, o que desliga a transcrição de um gene essencial.

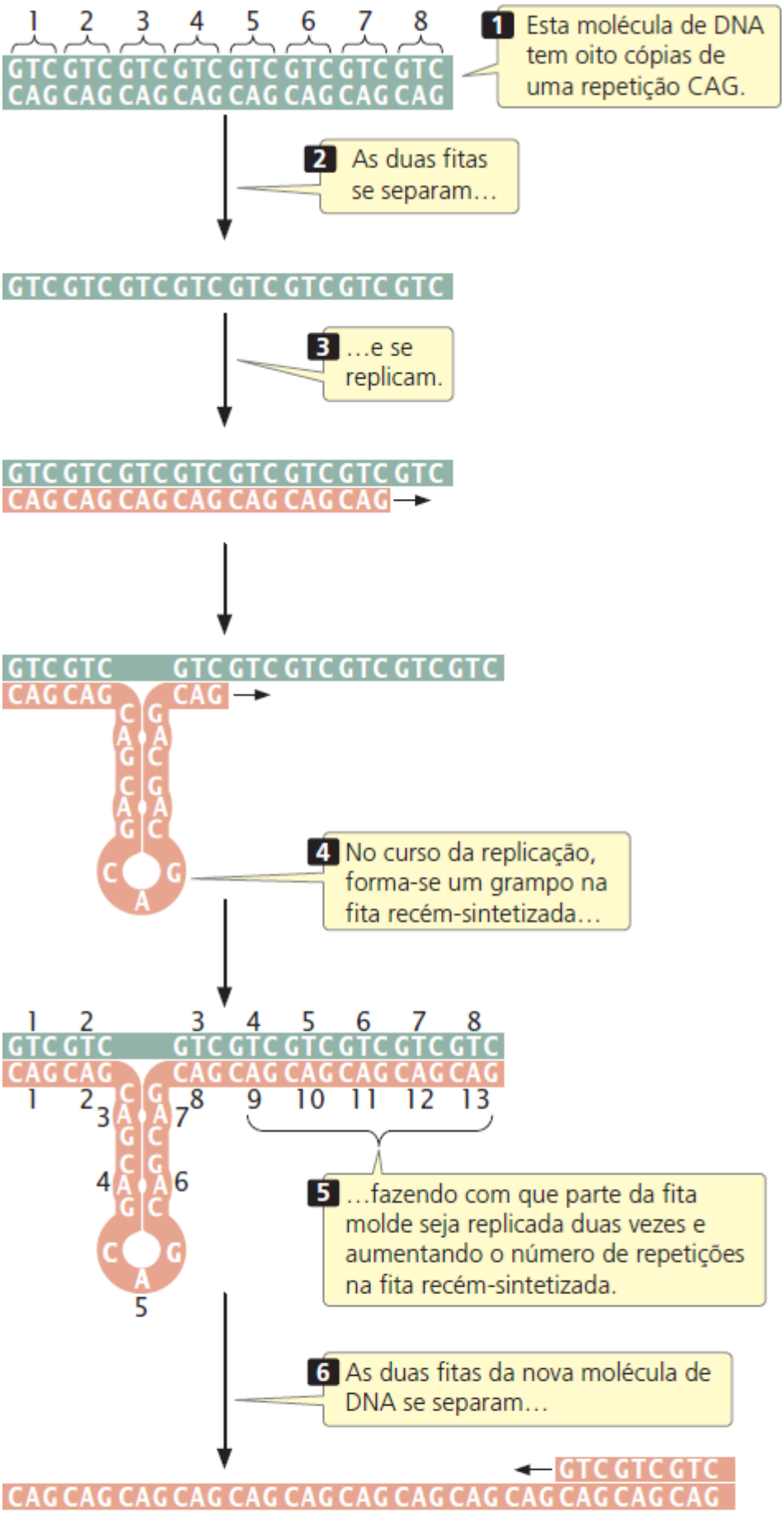
Existem algumas evidências sugerindo que a expansão das repetições de nucleotídios ocorre durante a replicação do DNA e parece estar relacionada com a formação de grampos (*hairpin*) e outras estruturas especiais de DNA que se formam no DNA de fita única composto por repetições de nucleotídios. Tais estruturas interferem na replicação normal ao provocar o deslize da fita, o mau alinhamento das sequências ou a interrupção da replicação. Um modelo de como os grampos de repetições poderiam resultar em expansão da repetição é apresentado na **Figura 18.4**.

Quadro 18.1 Exemplos de doenças genéticas humanas causadas por repetições expandidas de nucleotídios.

		Número de cópias de repetição	
Doença	**Sequência repetida**	**Faixa normal**	**Faixa da doença**
Atrofia muscular espinal e bulbar	CAG	11 a 33	40 a 62
Síndrome do X frágil	CGG	6 a 54	50 a 1.500
Síndrome de Jacobsen	CGG	11	100 a 1.000
Ataxia espinocerebelar (vários tipos)	CAG	4 a 44	21 a 130
Ataxia cerebelar autossômica dominante	CAG	7 a 19	37 a 220
Distrofia miotônica	CTG	5 a 37	44 a 3.000
Doença de Huntington	CAG	9 a 37	37 a 121
Ataxia de Friedrich	GAA	6 a 29	200 a 900
Atrofia dentato-rubra pálido-luisiana	CAG	7 a 25	49 a 75
Epilepsia mioclônica do tipo Unverricht-Lundborg	CCCCGCCCCGCG	2 a 3	12 a 13

Conceitos

As repetições expandidas de nucleotídios são regiões do DNA compostas de cópias repetidas de conjuntos de nucleotídios. Os números maiores de repetições de nucleotídios estão associados a várias doenças genéticas humanas.



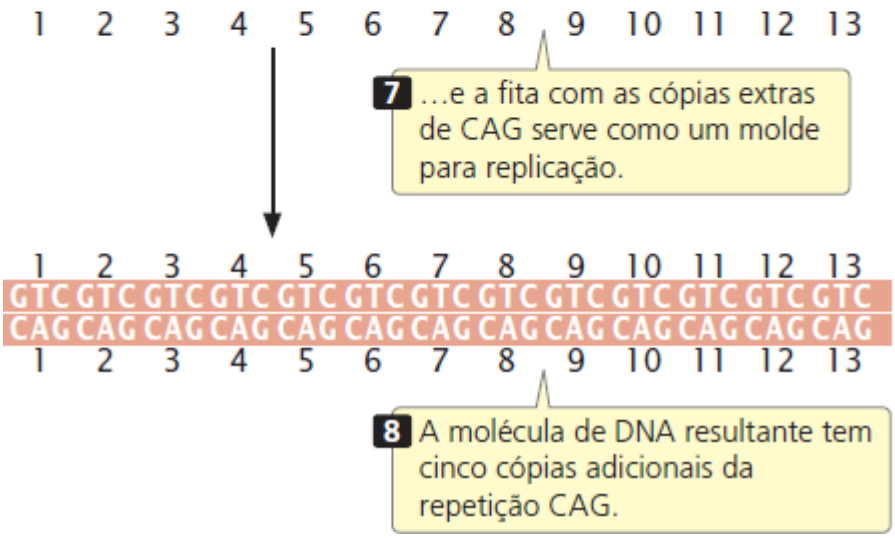


Figura 18.4 Um modelo de como o número de cópias de uma repetição de nucleotídios pode aumentar na replicação.

Efeitos fenotípicos das mutações

Outra forma usada para classificar as mutações são seus efeitos fenotípicos. No nível geral, podemos diferenciar uma mutação com base no seu fenótipo comparado com o fenótipo do tipo selvagem. Uma mutação que altera o fenótipo do tipo selvagem é chamada de **mutação direta**, enquanto uma **mutação reversa** (uma *reversão*) muda um genótipo mutante de volta para o tipo selvagem.

Os geneticistas usam outros termos para descrever os efeitos das mutações na estrutura da proteína. Uma substituição de base que resulta em um diferente aminoácido na proteína é chamada de **mutação *missense*** (**Figura 18.5 A**). Uma **mutação *nonsense*** muda um códon com sentido (que especifica um aminoácido) em um códon sem sentido (que termina a tradução), como apresentado na **Figura 18.5 B**. Se ocorrer uma mutação *nonsense* precoce na sequência de mRNA, a proteína será truncada e, em geral, não funcional.

Por causa da redundância do código genético, alguns códons diferentes especificam o mesmo aminoácido. Uma **mutação silenciosa** muda um códon para um códon sinônimo que especifica o mesmo aminoácido (**Figura 18.5 C**), alterando a sequência de DNA sem mudar a sequência de aminoácidos da proteína. Nem todas as mutações silenciosas, entretanto, são realmente silenciosas: algumas têm efeitos fenotípicos. Por exemplo, mutações silenciosas podem ter efeitos fenotípicos quando diferentes tRNAs (chamados de tRNA isoaceptores, ver *Capítulo 15*), são usados para diferentes códons sinônimos. Como alguns tRNAs isoaceptores são mais abundantes que outros, qual códon sinônimo é usado pode afetar a taxa de síntese de proteína. A taxa de síntese de proteína pode influenciar o fenótipo ao afetar a quantidade de proteína encontrada na célula e, em alguns casos, a dobra da proteína. Outras mutações silenciosas podem alterar sequências próximas de junções *éxon-íntron* que afetam a recomposição (ver *Capítulo 14*). Ainda, outras mutações silenciosas podem influenciar a ligação de miRNAs a sequências complementares no mRNA, que determinam se o mRNA será traduzido (ver *Capítulo 14*).

Uma **mutação neutra** é uma mutação *missense* que altera a sequência de aminoácidos da proteína mas não muda sua função de forma significativa. Mutações neutras ocorrem quando um aminoácido é substituído por outro que é quimicamente semelhante ou quando o aminoácido afetado tem pouca influência na função da proteína. Por exemplo, ocorrem mutações neutras em genes que codificam a hemoglobina; embora essas mutações alterem a sequência de aminoácidos da hemoglobina, elas não afetam sua capacidade de transportar o oxigênio.

As **mutações de perda de função** provocam perda completa ou parcial da função normal. Uma mutação de perda de função modifica a estrutura da proteína de modo que ela não funciona mais corretamente – ou a mutação pode ocorrer em regiões regulatórias que afetam a transcrição, a tradução ou a recomposição da proteína. Essas mutações são frequentemente recessivas: um organismo diploide tem de ser homozigoto para uma mutação de perda de função antes que os efeitos da perda da proteína funcional possam ser exibidos. As mutações que provocam a fibrose cística são do tipo com perda de função: elas produzem uma forma não funcional da proteína reguladora de condutância transmembrana da fibrose cística, que normalmente regula o movimento dos íons cloreto para dentro e para fora da célula (ver *Capítulo 5*).

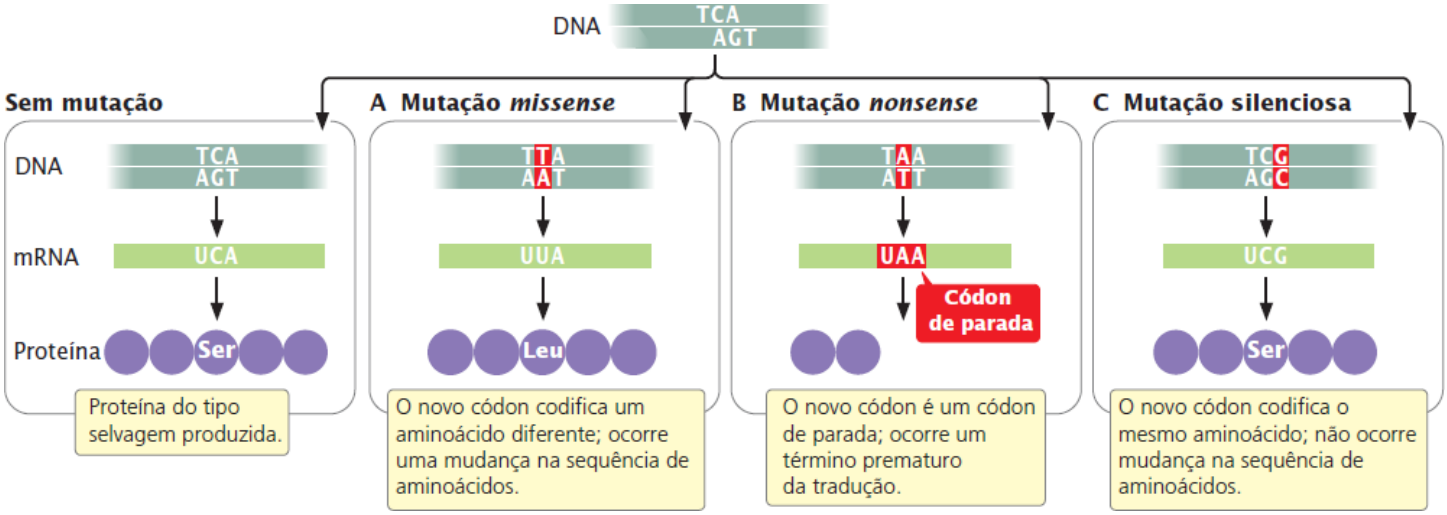


Figura 18.5 As substituições de base podem provocar mutações (A) *missense*, (B) *nonsense* e (C) silenciosas.

Por outro lado, uma **mutação de ganho de função** faz com que a célula produza uma proteína ou produto de gene cuja função não existe normalmente. Esse caso pode ser um produto totalmente novo do gene ou um produzido em um tecido inadequado ou em um momento inadequado do desenvolvimento. Por exemplo, uma mutação em um gene que codifica um fator de crescimento pode fazer com que o receptor mutante estimule o crescimento de forma constante, mesmo na ausência do fator de crescimento. As mutações de ganho de função são dominantes na sua expressão, porque uma única cópia da mutação leva a um novo produto gênico. Ainda, outros tipos de mutações são **mutações condicionais**, expressas apenas em algumas circunstâncias. Por exemplo, algumas mutações condicionais afetam o fenótipo apenas em temperaturas elevadas. Outras são **mutações letais**, que provocam a morte prematura (*Capítulo 5*). Resolva o Problema 22

Mutações supressoras

Uma **mutação supressora** é uma mudança genética que esconde ou suprime o efeito de outra mutação. Esse tipo de mutação é diferente de uma mutação reversa, na qual o sítio mutado volta para a sequência original do tipo selvagem (**Figura 18.6**). Uma mutação supressora ocorre em um sítio diferente do sítio da mutação original; assim, um indivíduo com mutação supressora é um duplo mutante, tendo a mutação original e a mutação supressora, mas exibindo o fenótipo de um tipo selvagem não mutante. Os geneticistas diferenciam entre duas classes de mutações supressoras: intragênicas e intergênicas.

Mutações supressoras intragênicas. Uma **mutação supressora intragênica** ocorre no mesmo gene que tem a mutação a ser suprimida e pode funcionar de várias maneiras. O supressor pode mudar um segundo nucleotídio no mesmo códon alterado pela mutação original, produzindo um códon que especifica o mesmo aminoácido que o especificado pelo códon original, não mutante (**Figura 18.7**). Os supressores intragênicos também podem suprimir uma mutação no quadro de leitura. Se a mutação original for a deleção de uma base, então a adição de uma única base em um local do gene restaura o quadro de leitura original.

Considere a sequência de nucleotídios a seguir na fita molde de DNA e os aminoácidos que ela codifica:

DNA	3′-AAA	TCA	CTT	GGC	GTA	CAA-5′
mRNA	5′-UUU	AGU	GAA	CCG	CAU	GUU-3′
Aminoácidos	Phe	Ser	Glu	Pro	His	Val

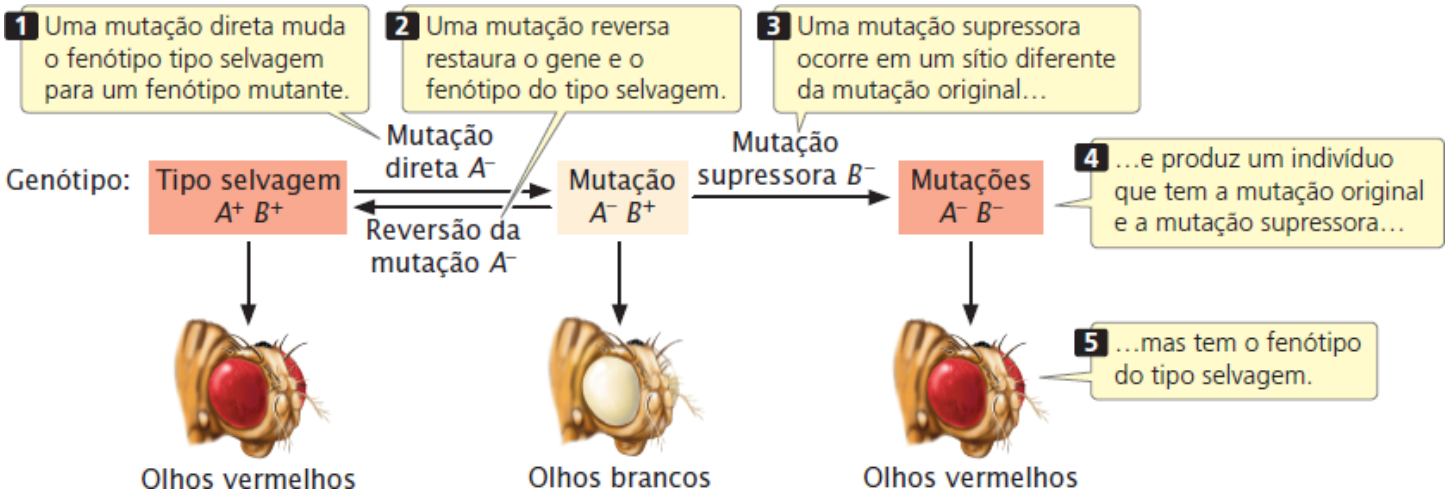


Figura 18.6 **Relação entre as mutações direta, reversa e supressora.**

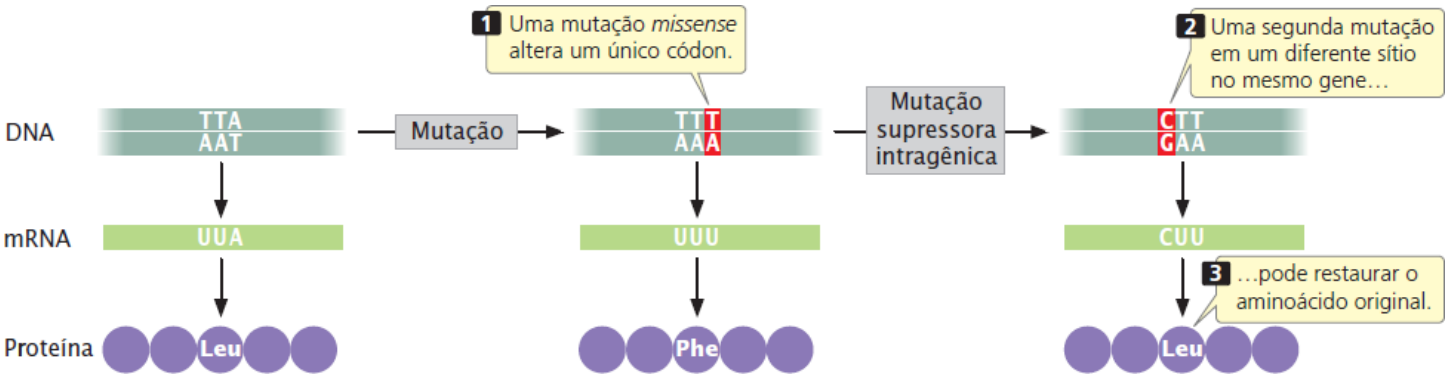


Figura 18.7 **Uma mutação supressora intragênica ocorre no gene que tem a mutação a ser suprimida.**

Suponha que ocorra a deleção de uma base no primeiro nucleotídio do segundo códon. Essa deleção muda o quadro de leitura em um nucleotídio e altera todos os aminoácidos que seguem a mutação.

Deleção de um nucleotídio

DNA	3′—AAA	XCAC	TTG	GCG	TAC	AA—5′
mRNA	5′—UUU	GUG	AAC	CGC	AUG	UU—3′
Aminoácidos	Phe	Val	Asn	Arg	Met	

Se um único nucleotídio for adicionado ao terceiro códon (a mutação supressora), o quadro de leitura é restaurado, embora dois dos aminoácidos sejam diferentes dos especificados pela sequência original.

Inserção de um nucleotídio

DNA	3′—AAA	CAC	TTT	GGC	GTA	CAA—5′
mRNA	5′—UUU	GUG	AAA	CCG	CAU	GUU—3′
Aminoácidos	Phe	Val	Lys	Pro	His	Val

Do mesmo modo, uma mutação por causa de uma inserção pode ser suprimida por uma subsequente deleção no mesmo gene.

Uma terceira maneira de um supressor intragênico atuar consiste em fazer mudanças compensatórias na proteína. Uma primeira mutação *missense* pode alterar a dobradura de uma cadeia de polipeptídios ao alterar a maneira como os aminoácidos na proteína interagem entre si. Uma segunda mutação *missense* em um diferente sítio (o supressor) pode recriar o padrão de

dobradura original ao restaurar as interações entre os aminoácidos.

Mutações supressoras intergênicas. Uma **mutação supressora intergênica**, ao contrário, ocorre em um gene diferente daquele que carreia a mutação original. Em alguns casos, esses supressores modificam a tradução do mRNA. Neste exemplo ilustrado na **Figura 18.8 A**, a sequência original de DNA é AAC (UUG no mRNA) e especifica a leucina. Esta sequência sofre mutação para ATC (UAG no mRNA), um códon de término (**Figura 18.8 B**). A mutação *nonsense* ATC poderia ser suprimida por uma segunda mutação em um gene diferente que codifica um tRNA; essa segunda mutação resultaria em um códon capaz de parear com o códon de término UAG (**Figura 18.8 C**). Por exemplo, o gene que codifica o tRNA para tirosina ($tRNA^{Tyr}$), que tem o anticódon AUA, sofreria mutação para ter o anticódon AUC, que então pareria com o códon de parada UAG. Em vez de a tradução terminar no códon UAG, a tirosina seria inserida na proteína e seria produzida uma proteína inteira, embora a tirosina agora substituísse a leucina. O efeito dessa mudança dependeria do papel desse aminoácido na estrutura geral da proteína, mas é provável que o efeito da mutação supressora seja menos prejudicial do que o efeito da mutação *nonsense*, que interrompe a tradução prematuramente.

Como as células em muitos organismos têm múltiplas cópias de genes do tRNA, outras cópias não mutantes de $tRNA^{Tyr}$ permaneceriam disponíveis para reconhecer os códons de tirosina nos transcritos do gene mutante em questão, e outros genes seriam expressos simultaneamente. Esperaríamos que os tRNAs que não sofreram a mutação supressora descrita também suprimiriam os códons de término normais nas extremidades das sequências codificadoras, resultando na produção de proteínas maiores que o normal, mas esse evento não ocorre.

Os supressores intergênicos também podem agir por interações gênicas (ver *Capítulo 5*). As cadeias de polipeptídios produzidas por dois genes podem interagir para produzir uma proteína funcional. Uma mutação em um gene pode alterar o polipeptídio codificado de modo que a interação dos dois polipeptídios seja destruída, e nesse caso não ocorre produção de uma proteína funcional. Uma mutação supressora no segundo gene pode produzir uma mudança compensatória no seu polipeptídio, portanto, restaurando a interação original. O **Quadro 18.2** apresenta as características de alguns tipos diferentes de mutações.

Conceitos

Uma mutação supressora cancela o efeito de uma mutação anterior em um sítio diferente. Uma mutação supressora intragênica ocorre no *mesmo* gene que tem a mutação original; uma mutação supressora intergênica ocorre em um gene *diferente*.

Checagem dos conceitos 2

Qual é a diferença entre mutação supressora e mutação reversa?

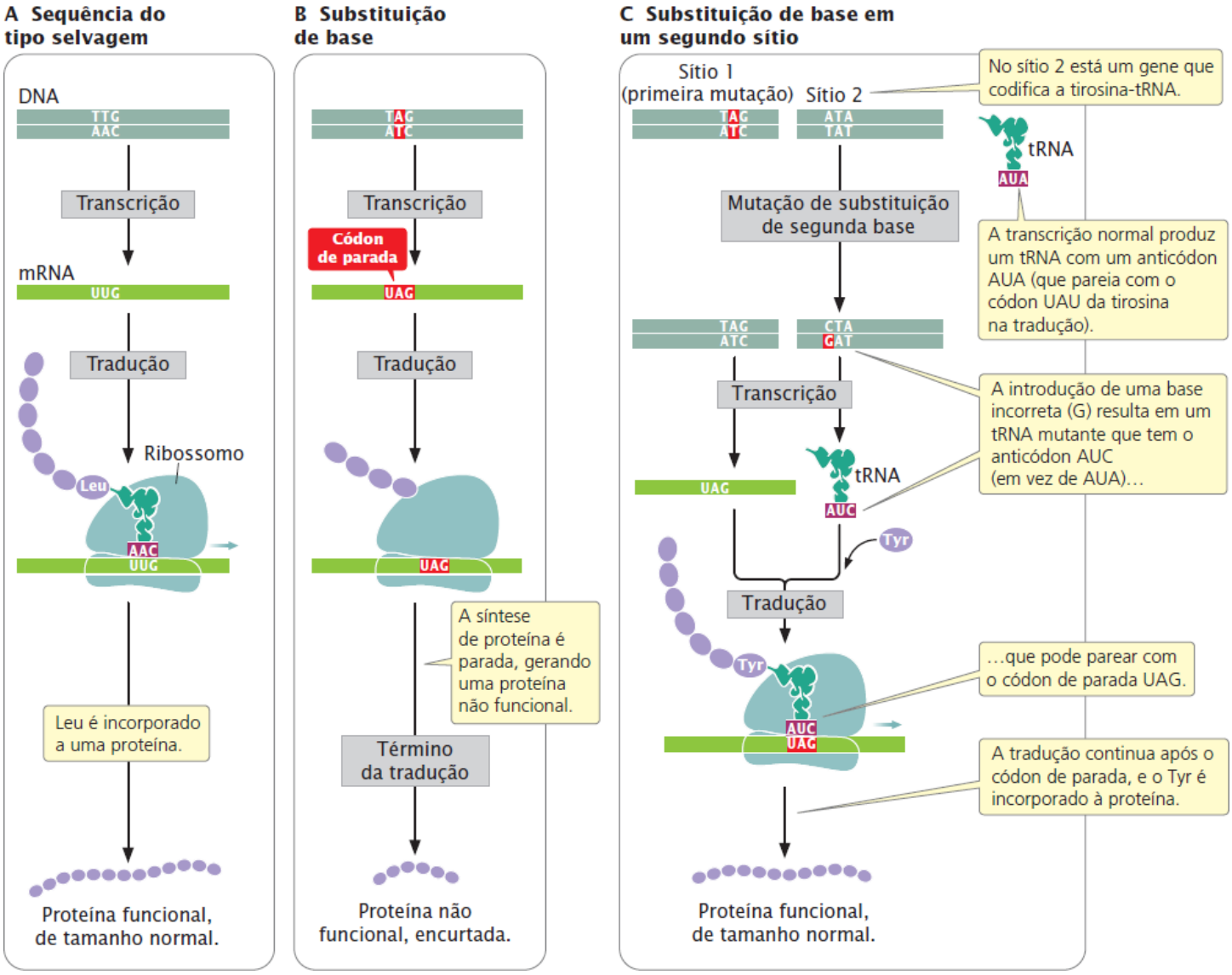


Figura 18.8 **A mutação supressora intergênica ocorre em um gene diferente daquele que carreia a mutação original. A.** A sequência do tipo selvagem produz uma proteína funcional, de tamanho normal. **B.** Uma substituição de base em um sítio no mesmo gene produz um códon de parada prematuro, resultando em uma proteína encurtada, não funcional. **C.** Uma substituição de base em um sítio em outro gene, que nesse caso codifica tRNA, altera o anticódon do $tRNA^{Tyr}$; o $tRNA^{Tyr}$ pode parear com o códon de parada produzido pela mutação original, fazendo com que a tirosina seja incorporada à proteína e a tradução continue.

Problema desenvolvido

Um gene codifica a proteína com a seguinte sequência de aminoácidos:

Met-Arg-Cys-Ile-Lys-Arg

Uma mutação de um único nucleotídio altera a sequência de aminoácidos para:

Met-Asp-Ala-Tyr-Lys-Gly-Glu-Ala-Pro-Val

Uma segunda mutação de único nucleotídio ocorre no mesmo gene e suprime os efeitos da primeira mutação (um supressor intragênico). Com a mutação original e o supressor intragênico, a proteína tem a seguinte sequência de aminoácidos:

Met-Asp-Gly-Ile-Lys-Arg

Qual é a natureza e a localização da primeira mutação e da mutação supressora intragênica?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

O tipo e a localização da primeira mutação e da mutação supressora intragênica.

Quais informações são fornecidas para solucionar o problema?

- A sequência de aminoácidos da proteína codificada pelo gene original não mutante
- A sequência de aminoácidos da proteína codificada pelo gene mutante
- A sequência de aminoácidos da proteína codificada pelo gene mutante e o supressor intragênico.

Etapas para a solução

A primeira mutação modifica o quadro de leitura, porque todos os aminoácidos após o Met são alterados. As inserções e deleções afetam o quadro de leitura; a mutação original é composta de inserção ou deleção de único nucleotídio no segundo códon. O supressor intragênico restaura o quadro de leitura; é mais provável que o supressor intragênico também seja uma inserção ou deleção de único nucleotídio. Se a primeira mutação for uma inserção, o supressor tem de ser uma deleção; se a primeira mutação for uma deleção, então o supressor tem de ser uma inserção. Observe que a proteína produzida pelo supressor ainda é diferente da proteína original no segundo e no terceiro aminoácido, mas o segundo aminoácido produzido pelo supressor é o mesmo que na proteína produzida pela mutação original. Assim, a mutação supressora ocorreu obrigatoriamente no terceiro códon, porque o supressor não altera o segundo aminoácido.

Quadro 18.2 Características dos diferentes tipos de mutações.

Tipo de mutação	Definição
Substituição de base	Muda a base de um único nucleotídio de DNA
Transição	Substituição de base na qual uma purina substitui uma purina ou uma pirimidina substitui uma pirimidina
Transversão	Substituição de base na qual uma purina substitui uma pirimidina ou uma pirimidina substitui uma purina
Inserção	Adição de um ou mais nucleotídios
Deleção	Deleção de um ou mais nucleotídios
Mutação no quadro de leitura	Inserção ou deleção que altera o quadro de leitura de um gene
Deleção ou inserção no quadro	Deleção ou inserção de um múltiplo de três nucleotídios que não altera o quadro de leitura
Repetições expandidas de nucleotídios	Sequência repetida de um conjunto de nucleotídios no qual o número de cópias de sequência aumenta
Mutação direta	Muda o fenótipo do tipo selvagem para um fenótipo mutante
Mutação reversa	Muda uma mutação do fenótipo de volta para o fenótipo do tipo selvagem
Mutação *missense*	Muda um códon com sentido em um diferente códon com sentido, resultando na incorporação de um aminoácido diferente na proteína
Mutação *nonsense*	Muda um códon com sentido em um códon sem sentido (de parada), levando ao término prematuro da tradução
Mutação silenciosa	Muda um códon com sentido em um códon sinônimo,

	deixando a sequência de aminoácidos da proteína inalterada
Mutação neutra	Muda a sequência de aminoácidos de uma proteína sem alterar sua capacidade de funcionar
Mutação de perda de função	Leva à perda completa ou parcial da função
Mutação de ganho de função	Leva ao surgimento de um novo traço ou função ou leva ao surgimento de um traço em um tecido inadequado ou em um momento inadequado
Mutação letal	Leva à morte prematura
Mutação supressora	Suprime o efeito de uma mutação anterior em um sítio diferente
Mutação supressora intragênica	Suprime o efeito de uma mutação anterior no mesmo gene
Mutação supressora intergênica	Suprime o efeito de uma mutação anterior em outro gene

Solucione o **Problema 23** no final deste capítulo para praticar mais sobre a análise de mutações.

Taxas de mutação

A frequência com a qual o alelo do tipo selvagem em um *locus* muda para um alelo mutante é chamada de **taxa de mutação** e em geral é expressa como o número de mutações por unidade biológica, que podem ser mutações por divisão celular, por gameta ou por ciclo de replicação. Por exemplo, a acondroplasia é um tipo de nanismo hereditário nos humanos resultante de uma mutação dominante. Em geral, surgem cerca de quatro acondroplasias em cada 100.000 gametas, então a taxa de mutação é 4/100.000 ou 0,00004 mutação por gameta. A taxa de mutação fornece informações sobre a frequência de uma mutação.

Fatores que afetam as taxas de mutação. Os cálculos das taxas de mutação são afetados por três fatores. Primeiro, eles dependem da frequência com que uma mudança ocorre no DNA. Uma mudança no DNA pode surgir a partir de mudanças moleculares espontâneas no DNA ou pode ser induzida por agentes químicos, biológicos ou físicos no ambiente.

O segundo fator que influencia a taxa de mutação é a probabilidade de reparo quando uma mudança ocorrer. A maioria das células tem vários mecanismos para reparar o DNA alterado, então a maioria das alterações é corrigida antes de ser replicada. Se esses sistemas de reparo forem eficientes, as taxas de mutação serão baixas; se eles forem defeituosos, as taxas de mutação serão elevadas. Algumas mutações aumentam a taxa geral de mutação em outros genes; elas ocorrem em geral em genes que codificam componentes do maquinário de replicação ou em enzimas de reparo do DNA.

O terceiro fator é a probabilidade de que uma mutação seja detectada. Quando o DNA é sequenciado, todas as mutações são potencialmente detectáveis. Na prática, entretanto, as mutações são detectadas por meio dos seus efeitos fenotípicos. Algumas mutações podem surgir em uma taxa maior simplesmente porque é mais fácil detectá-las.

Variação nas taxas de mutação. Podemos tirar várias conclusões gerais sobre as taxas de mutação, embora elas variem entre os genes e as espécies (**Quadro 18.3**). Primeiro, as taxas de mutação espontânea são baixas para todos os organismos estudados. As típicas taxas de mutação para os genes bacterianos são cerca de 1 para 100 mutações por 10 bilhões de células (de 1×10^{-8} a 1×10^{-10}). As típicas taxas de mutação para os genes eucarióticos são um pouco maiores, cerca de 1 para 10 mutações por milhão de gametas (de 1×10^{-5} a 1×10^{-6}). Esses valores maiores nos eucariotos podem ocorrer porque as taxas são calculadas por gameta e são necessárias várias divisões celulares para produzir um gameta, enquanto as taxas de mutação nas células procarióticas são calculadas *por divisão celular*.

Quadro 18.3 Taxas de mutação de diferentes genes em diferentes organismos.

Organismo	Mutação	Taxa	Unidade

Bacteriófago T2	Inibição da lise	1×10^{-8}	Por replicação
	Gama de hospedeiros	3×10^{-9}	
Escherichia coli	Fermentação da lactose	2×10^{-7}	Por divisão celular
	Necessidade de histidina	2×10^{-8}	
Neurospora crassa	Necessidade de inositol	8×10^{-8}	Por esporo assexuado
	Necessidade de adenina	4×10^{-8}	
Milho	Cor do grão	$2{,}2 \times 10^{-6}$	Por gameta
Drosophila	Cor dos olhos	4×10^{-5}	Por gameta
	Aloenzimas	$5{,}14 \times 10^{-6}$	
Camundongo	Cor do pelo do albino	$4{,}5 \times 10^{-5}$	Por gameta
	Cor da diluição do pelo	3×10^{-5}	
Humano	Doença de Huntington	1×10^{-6}	Por gameta
	Acondroplasia	1×10^{-5}	
	Neurofibromatose (Michigan)	1×10^{-4}	
	Hemofilia A (Finlândia)	$3{,}2 \times 10^{-5}$	
	Distrofia muscular de Duchenne (Wisconsin)	$9{,}2 \times 10^{-5}$	

As diferenças nas taxas de mutação entre as espécies podem ocorrer graças a diferentes capacidades para reparar as mutações, exposições desiguais a mutágenos ou diferenças biológicas nas taxas de mutações que surgem espontaneamente. Mesmo em uma única espécie, as taxas espontâneas de mutação variam entre os genes. O motivo para essa variação ainda não está totalmente compreendido, mas algumas regiões do DNA são conhecidas como pontos de acesso para as mutações.

Pesquisa recente sugere que menos mutações ocorrem em sequências de DNA ocupadas pelos nucleossomos (ver *Capítulo 1*). Podem ocorrer taxas de mutação reduzidas nessas sequências porque o DNA associado aos nucleossomos está menos exposto a mutágenos, mas também poderia ser explicado pelo efeito dos nucleossomos sobre o reparo, a recombinação ou a replicação do DNA, visto que todos esses fenômenos influenciam a taxa de mutação.

Vários estudos recentes mediram as taxas de mutação diretamente ao sequenciar os genes de organismos antes e após várias gerações. Esses novos estudos sugerem que as taxas de mutação são maiores que as previamente medidas com base nas mudanças nos fenótipos. Em um estudo, os geneticistas sequenciaram segmentos aleatoriamente escolhidos de DNA no nematódeo *Caenorhabditis elegans* e encontraram cerca de 2,1 mutações por genoma por geração, que era 10 vezes maior que as estimativas anteriores com base nas mudanças fenotípicas. Os pesquisadores descobriram que cerca da metade das mutações eram inserções e deleções.

O recente sequenciamento de genoma também forneceu informações mais acuradas sobre as taxas de mutação nos humanos. Vários estudos de sequenciamento sugerem que a taxa geral de substituições de base nos humanos é cerca de 1×10^{-8} mutações por par de bases por geração. Outra pesquisa sugere que cada pessoa carreia aproximadamente 100 mutações de perda de função na linhagem germinativa.

Mutação adaptativa. Como vamos estudar nos Capítulos 24 a 26, a variação genética é crítica para a mudança evolutiva que produz adaptação a novos ambientes. Novas variantes genéticas surgem principalmente por meio da mutação. Por muitos anos, acreditava-se que a variação genética surgisse de forma aleatória e em taxas que eram independentes da necessidade de adaptação. Entretanto, evidências sugerem que ambientes estressantes – nos quais a adaptação pode ser necessária para sobreviver – podem induzir mais mutações nas bactérias, um processo chamado de **mutação adaptativa**. A ideia de mutação adaptativa já foi muito discutida; os críticos questionam que a maioria das mutações seria prejudicial e então uma maior mutagênese seria danosa para a maioria dos casos.

Conceitos

A taxa de mutação é a frequência de ocorrência de uma mutação específica. As taxas de mutação em geral são baixas e são afetadas por fatores ambientais e genéticos.

Checagem dos conceitos 3

Quais são os três fatores que afetam as taxas de mutação?

18.2 As mutações são potencialmente causadas por vários fatores diferentes

As mutações resultam tanto de fatores internos quanto externos. As que surgem sob condições normais são chamadas de **mutações espontâneas**, enquanto as que surgem a partir de mudanças causadas por substâncias químicas ou radiação do ambiente são chamadas de **mutações induzidas**.

Erros espontâneos de replicação

A replicação é incrivelmente acurada: surge menos de um erro em um bilhão de nucleotídios durante a síntese do DNA (ver *Capítulo 12*). Entretanto, ocorrem erros espontâneos de replicação em alguns casos.

Shifts tautoméricos. Antigamente, acreditava-se que a causa primária dos erros espontâneos de replicação fossem os *shifts* tautoméricos, nos quais as posições dos prótons nas bases de DNA mudam. As bases purina e pirimidina existem em formas químicas diferentes, chamadas tautômeros. As duas formas tautoméricas de cada base estão em equilíbrio dinâmico, embora uma forma seja mais comum que a outra. Os pareamentos padrão de base de Watson e Crick – adenina com timina e citosina com guanina – ocorrem entre as formas comuns das bases, mas se as bases estiverem em suas raras formas tautoméricas, outros pareamentos de base são possíveis. Por exemplo, a forma comum de citosina pareia com guanina, mas o tautômero raro de citosina pareia com adenina.

Watson e Crick propuseram que os *shifts* tautoméricos poderiam provocar mutações, e por muitos anos essa proposta foi o modelo aceito para os erros espontâneos de replicação. Entretanto, nunca houve evidências convincentes de que os raros tautômeros são a causa das mutações espontâneas. Além disso, a pesquisa atualmente mostra poucas evidências de tautômeros no DNA.

Pareamento incorreto causado por outras estruturas. Os pareamentos incorretos surgem frequentemente por oscilação (*wobble*) (ver *Capítulo 15*), em que as formas normais, protonadas, e outras formas de bases são capazes de parear graças à flexibilidade na estrutura helicoidal do DNA (**Figura 18.9**). Essas estruturas foram detectadas em moléculas de DNA e atualmente acredita-se que sejam responsáveis por muitos dos pareamentos incorretos na replicação.

Erros de incorporação e erros de replicação. Quando uma base malpareada é incorporada a uma cadeia de nucleotídios recém-sintetizada, é dito que um ocorreu um **erro de incorporação**. Suponha que, na replicação, a timina (que normalmente pareia com adenina) faça um pareamento incorreto com guanina por causa de oscilação (**Figura 18.10**). No próximo ciclo de replicação, as duas bases malpareadas se separam e cada uma serve como molde para a síntese de uma nova fita de nucleotídios. Nesse caso, a timina pareia com adenina, produzindo outra cópia da sequência original do DNA. Por outro lado, entretanto, a guanina incorretamente incorporada serve como molde e pareia com citosina, produzindo uma nova molécula de DNA que tem um erro – um par C • G no lugar do par T • A original (uma substituição de base T • A → C • G). O erro original de incorporação leva a um **erro de replicação**, o que cria uma mutação permanente porque todos os pareamentos de base estão corretos e não existe mecanismo para sistemas de reparo para detectar o erro.

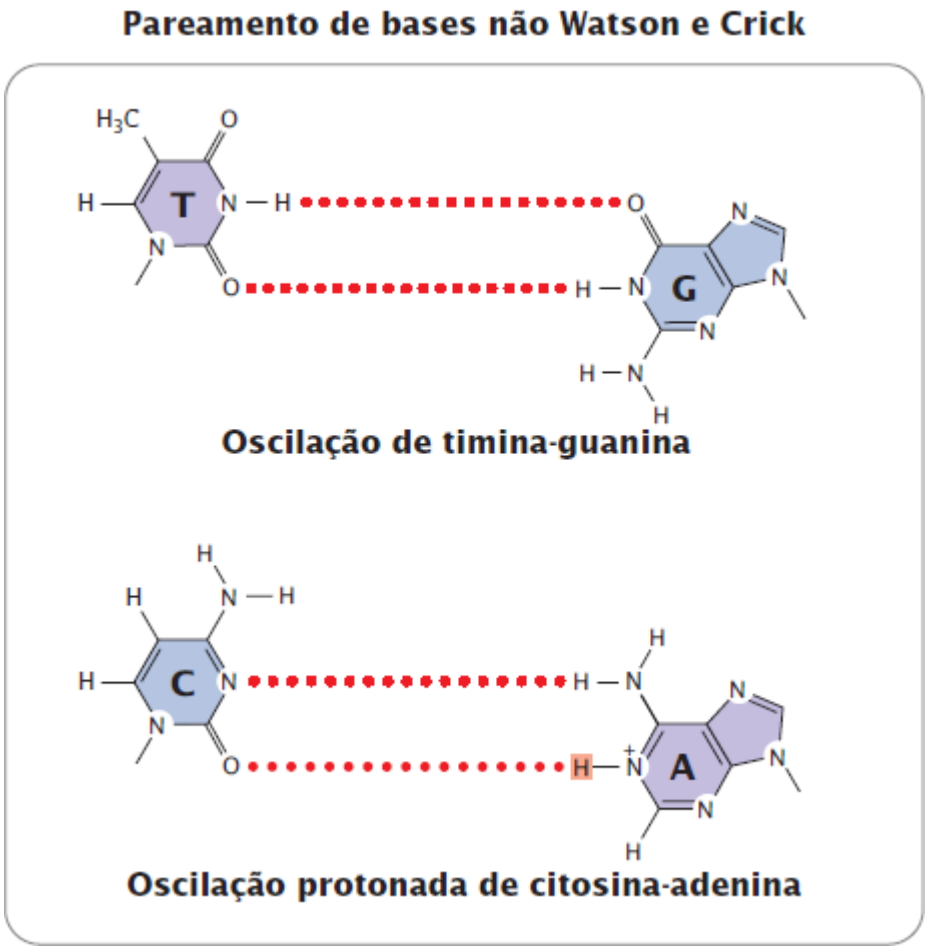


Figura 18.9 Os pareamentos de base não padrões podem surgir como resultado da flexibilidade na estrutura do DNA. A timina e a guanina podem parear por oscilação entre as bases normais. No caso da citosina e da adenina, elas podem parear por oscilação quando a adenina está protonada (tem um átomo de hidrogênio extra).

Causas das deleções e inserções. As mutações causadas por pequenas inserções e deleções também surgem espontaneamente na replicação e no *crossing over*. Pode ocorrer o **deslizamento da fita** quando uma fita de nucleotídios forma uma pequena alça (**Figura 18.11**). Se os nucleotídios desenrolados estiverem na fita recém-sintetizada, surgirá uma inserção. No próximo ciclo de replicação, a inserção será replicada e ambas as fitas terão a inserção. Se os nucleotídios desenrolados estiverem na fita molde, então a fita recém-replicada terá uma deleção e essa deleção será perpetuada em ciclos subsequentes de replicação.

Outro processo que produz inserções e deleções é o *crossing over* desigual. No *crossing over* normal, as sequências homólogas das duas moléculas de DNA se alinham e o *crossing over* não altera o número de nucleotídios nas moléculas. O pareamento desalinhado pode provocar o ***crossing over* desigual**, que resulta em uma molécula de DNA com uma inserção e a outra com uma deleção **(Figura 18.12)**.

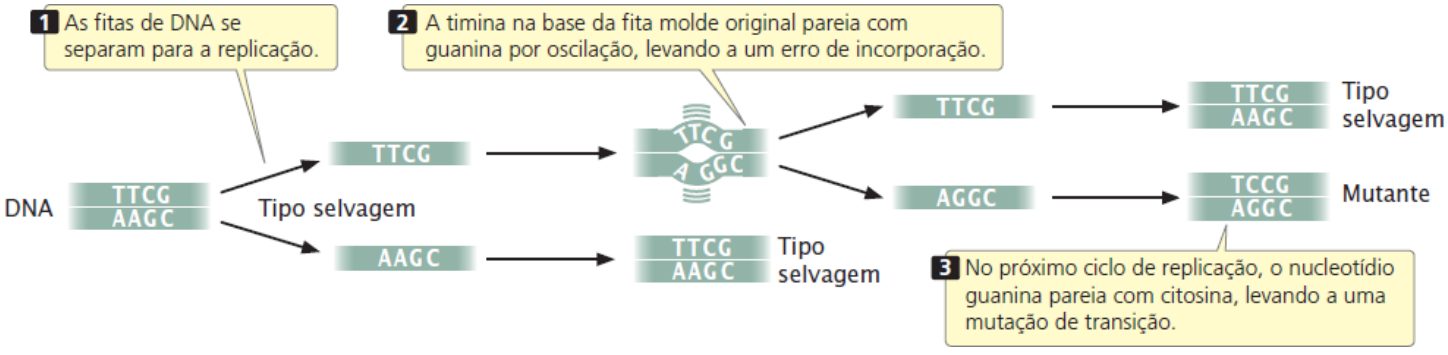


Figura 18.10 O pareamento de base por oscilação leva a um erro de replicação.

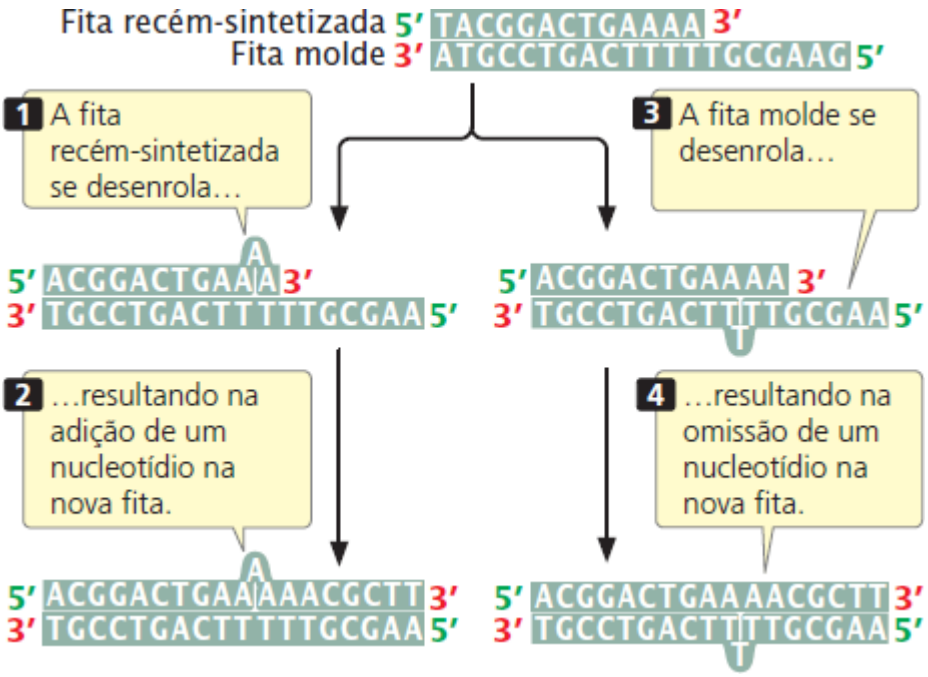


Figura 18.11 **As inserções e deleções podem surgir a partir do deslizamento da fita.**

Conceitos

Os erros espontâneos de replicação surgem a partir de estruturas com base alterada e pareamento de base em oscilação. Podem ocorrer pequenas inserções e deleções por meio do deslizamento da fita na replicação e do *crossing over* desigual.

Mudanças químicas espontâneas

Além das mutações espontâneas que surgem na replicação, as mutações também são resultado de mudanças químicas espontâneas no DNA. Uma dessas mudanças é a **depurinação**, a perda de uma base purina de um nucleotídio. A depurinação surge quando a ligação covalente que conecta a purina ao átomo carbono 1′ do açúcar desoxirribose se rompe (**Figura 18.13 A**), produzindo um sítio apurínico, um nucleotídio sem sua base purina. Um sítio apurínico não serve como molde para uma base complementar na replicação. Se não houver restrições de pareamento de base, um nucleotídio incorreto (na maioria das vezes a adenina) é incorporado à fita de DNA recém-sintetizada oposta ao sítio apurínico (**Figura 18.13 B**), frequentemente levando a um erro de incorporação. Esse erro é então transformado em um erro de replicação no próximo ciclo de replicação. A depurinação é a causa mais comum de mutação espontânea; uma célula de mamífero em cultura perde aproximadamente 10.000 purinas todos os dias. Também ocorre a perda das bases pirimidina, mas em uma taxa muito menor que a depurinação.

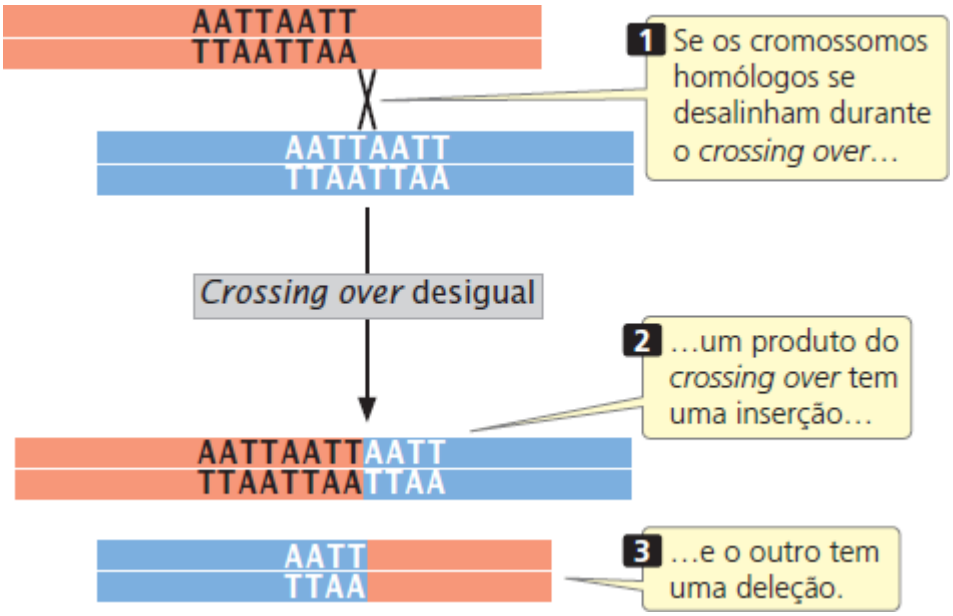


Figura 18.12 **O *crossing over* desigual produz inserções e deleções.**

Outra mudança química de ocorrência espontânea no DNA é a **desaminação**, a perda de um grupo amino (NH_2) de uma base. A desaminação pode ser espontânea ou induzida por substâncias químicas mutagênicas.

Esse fenômeno pode alterar as propriedades de pareamento de uma base: a desaminação da citosina, por exemplo, produz uracila (**Figura 18.14 A**) que pareia com adenina na replicação. Após outro ciclo de replicação, a adenina vai parear com timina, criando um par T • A no lugar do par original C • G (C • G → U • A → T • A); essa mudança química é uma mutação de transição. Esse tipo de mutação é evitado por enzimas que removem a uracila sempre que ela é encontrada no DNA. A capacidade de reconhecer o produto da desaminação da citosina pode explicar por que a timina, e não a uracila, é encontrada no DNA. Nos mamíferos, incluindo humanos, algumas bases de citosina no DNA estão naturalmente metiladas e existem na forma de 5-metilcitosina (5 mC, ver Figura 10.19). Quando desaminada, 5 mC se torna timina (**Figura 18.14 B**). Como a timina pareia com adenina na replicação, a desaminação da 5-metilcitosina muda um par original C • G para T • A (C • G → 5 mC • G → T • G → T • A). Consequentemente, as transições C • G → T • A são frequentes nas células dos mamíferos, e os sítios 5 mC são sítios quentes de mutação nos humanos.

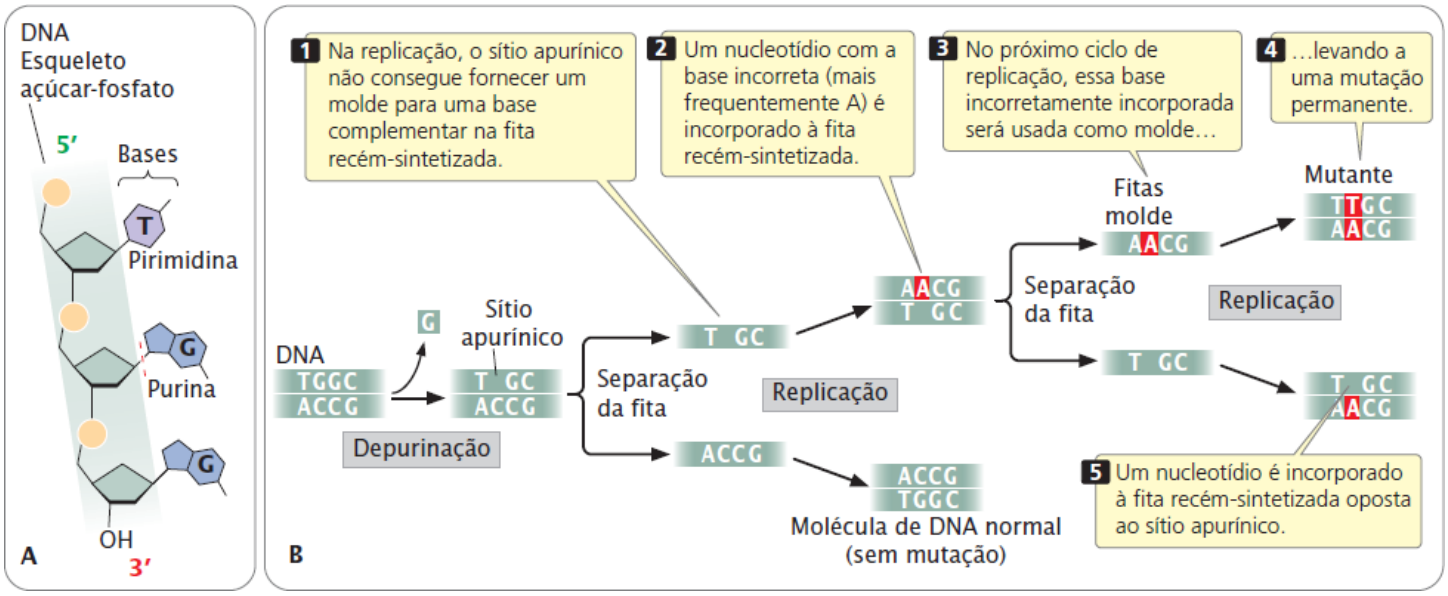


Figura 18.13 **A depurinação (perda de uma base purina de um nucleotídio) produz um sítio apurínico.**

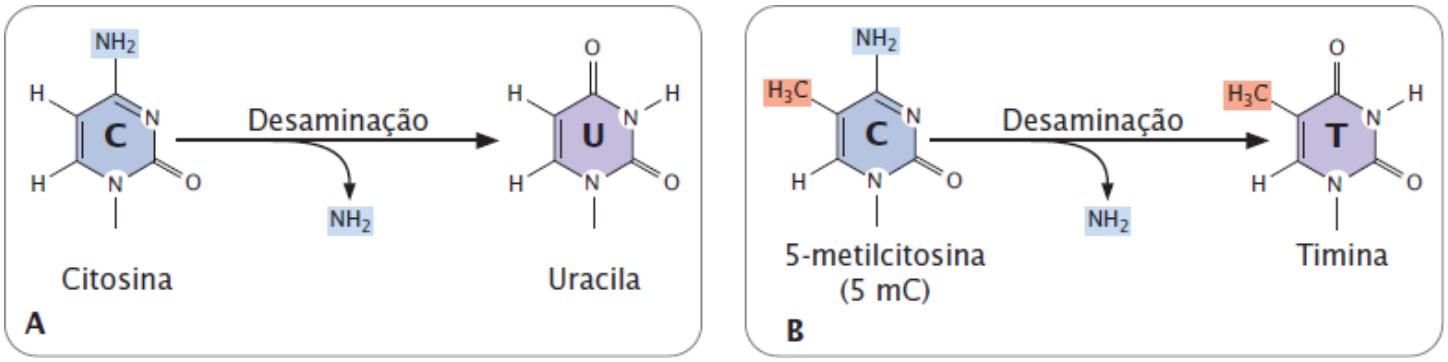


Figura 18.14 **A desaminação altera as bases do DNA.** Resolva o Problema 27

Conceitos

Algumas mutações surgem a partir de alterações espontâneas na estrutura do DNA, como a depurinação e a desaminação, que podem alterar as propriedades de pareamento das bases e provocam erros nos ciclos subsequentes de replicação.

Mutações quimicamente induzidas

Embora muitas mutações surjam espontaneamente, vários agentes ambientais são capazes de danificar o DNA, incluindo algumas substâncias químicas e radiação. Qualquer agente ambiental que aumenta de forma significativa a taxa de mutação

acima da taxa espontânea é chamado de **mutágeno**.

A primeira descoberta de um mutágeno químico foi feita por Charlotte Auerbach, que iniciou sua carreira em Berlim pesquisando o desenvolvimento de mutantes na *Drosophila*. Ao ter que lidar com o crescente antissemitismo na Alemanha nazista, Auerbach emigrou para a Grã-Bretanha em 1933. Ela continuou sua pesquisa sobre a *Drosophila*, e em 1940 começou a colaborar com o farmacologista John Robson sobre os efeitos mutagênicos do gás mostarda, que foi usado como arma química na Primeira Guerra Mundial. As condições experimentais eram rudimentares. Eles aqueciam o gás mostarda líquido em um bico de Bunsen no teto do prédio de farmacologia e as moscas eram expostas ao gás em uma grande câmara. Após desenvolver graves queimaduras em suas mãos com o gás, Auerbach deixou os outros colaboradores realizarem a exposição e ela analisou as moscas. Auerbach e Robson mostraram que o gás mostarda é de fato um poderoso mutágeno, reduzindo a viabilidade dos gametas e aumentando os números de mutações observados nos descendentes das moscas expostas. Como a pesquisa era parte de um esforço de segredo de guerra, a publicação de seus achados foi postergada até 1947.

Análogos da base. Uma classe de mutágenos químicos é composta de **análogos da base**, substâncias químicas com estruturas semelhantes à de qualquer uma das quatro bases padrão do DNA. As DNA polimerases não conseguem diferenciar entre os análogos das bases padrão; então, se esses análogos forem encontrados durante a replicação, eles podem ser incorporados a moléculas de DNA recém-sintetizadas. Por exemplo, 5-bromouracila (5BU) é o análogo da timina; ele tem a mesma estrutura da timina, exceto por ter um átomo de bromo (Br) no átomo do carbono 5 em vez de um grupo metila (**Figura 18.15 A**). Normalmente, a 5-bromouracila pareia com adenina como a timina faz, mas em alguns casos, pareia de forma incorreta com guanina (**Figura 18.15 B**), levando à transição (T • A → 5BU • A → 5BU • G → C • G), como apresentada na **Figura 18.16**. Por meio do pareamento incorreto, a 5-bromouracila também pode ser incorporada a uma fita de DNA recém-sintetizada oposta à guanina. No próximo ciclo de replicação, a 5-bromouracila pareia com a adenina, levando a outra transição (G • C → G • 5BU → A • 5BU → A • T).

No laboratório, as mutações causadas pelos análogos de bases podem ser revertidas ao tratar com o mesmo análogo ou com um análogo diferente.

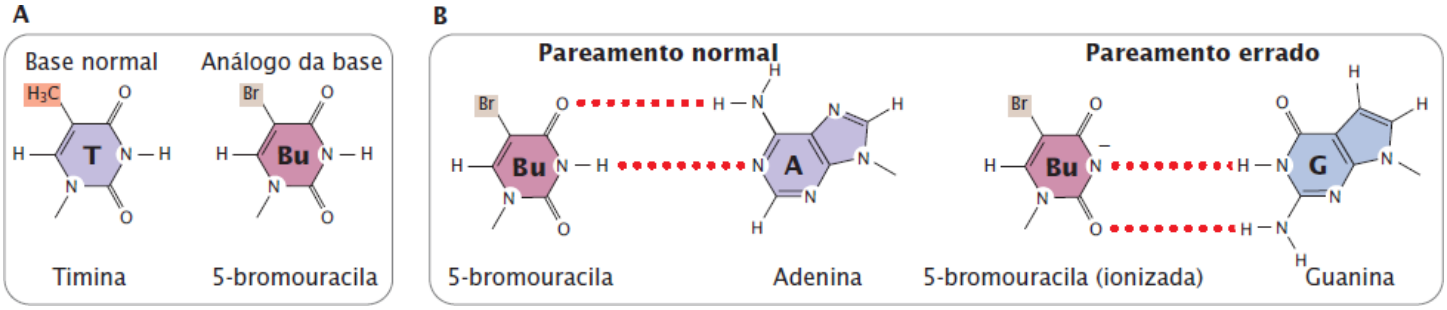


Figura 18.15 A 5-bromouracila (um análogo de base) é semelhante à timina, exceto por ter um átomo de bromo no lugar do grupo metila no átomo do carbono 5. Por causa da semelhança de suas estruturas, a 5-bromouracila pode ser incorporada ao DNA no lugar da timina. Semelhante à timina, a 5-bromouracila pareia normalmente com a adenina, mas quando ionizada, pode parear com guanina por oscilação.

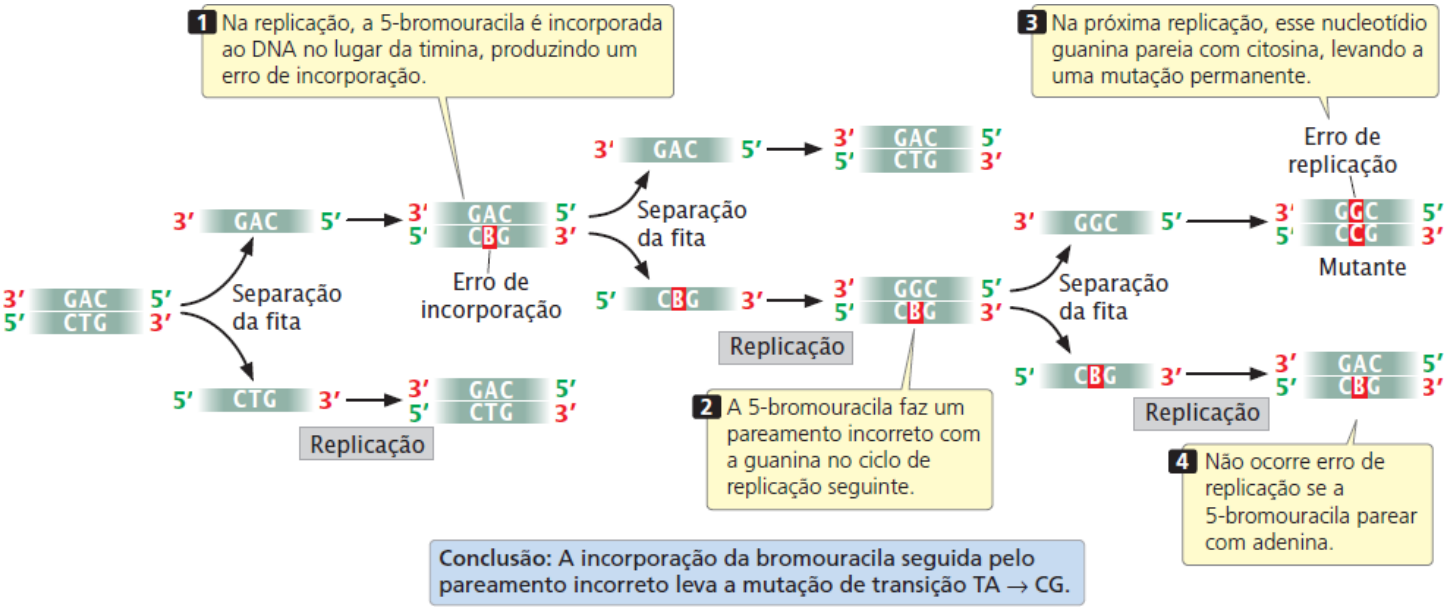


Figura 18.16 A 5-bromouracila pode levar a um erro de replicação.

Agentes alquilantes. Os agentes alquilantes são substâncias químicas que doam grupos alquila, como metila (CH_3) e etila (CH_3-CH_2) a bases nucleotídios. Por exemplo, etilmetilsulfonato (EMS) adiciona um grupo etila à guanina, produzindo O^6-etilguanina, que pareia com timina (**Figura 18.17 A**). Assim, o EMS produz transições C • G → T • A. O EMS também é capaz de adicionar um grupo etila à timina, produzindo 4-etiltimina, que então pareia com guanina, levando à transição T • A → C • G. Como o EMS produz as transições C • G → T • A e T • A → C • G, as mutações produzidas por esse agente podem ser revertidas com o tratamento adicional desse mesmo agente. O gás mostarda é outro agente alquilante.

Desaminação. Além da sua ocorrência espontânea (ver Figura 18.14), a desaminação pode ser induzida por algumas substâncias químicas. Por exemplo, o ácido nitroso desamina a citosina, criando uracila, que no próximo ciclo de replicação pareia com adenina (**Figura 18.17 B**), produzindo uma mutação de transição C • G → T • A. O ácido nitroso altera adenina para hipoxantina, que pareia com citosina levando a uma transição T • A → C • G. Ele também desamina guanina, produzindo xantina, que pareia com citosina como a guanina faz; entretanto, a xantina também pode parear com timina, gerando uma transição C • G → T • A. Ele produz exclusivamente mutações de transição, e como ambas as transições C • G → T • A e T • A → C • G são produzidas, essas mutações podem ser revertidas com ácido nitroso.

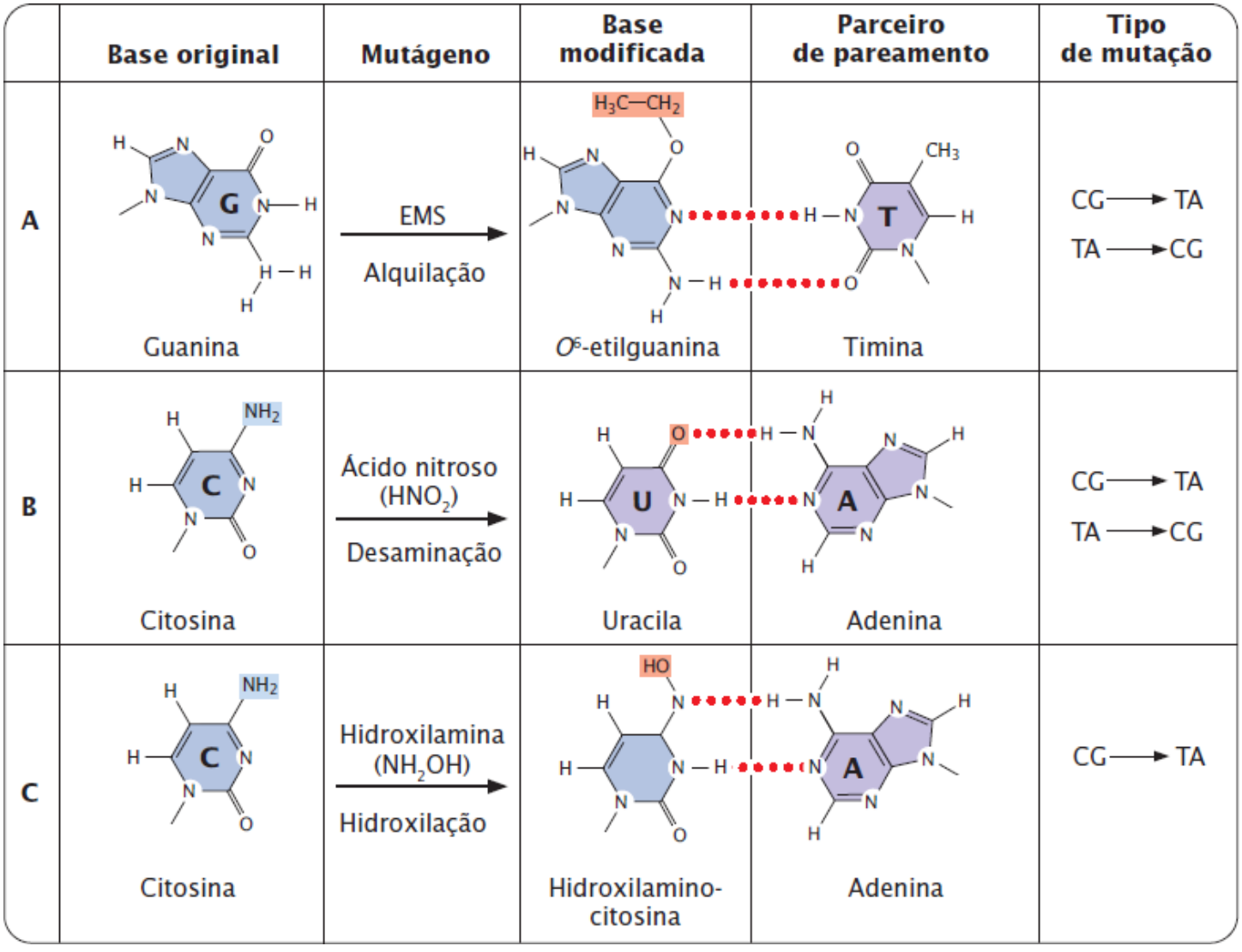


Figura 18.17 As substâncias químicas podem alterar as bases de DNA. Este quadro mostra alguns exemplos de mutações produzidas por agentes químicos.

Hidroxilamina. A hidroxilamina é um mutágeno modificador de base muito específico que adiciona um grupo hidroxila à citosina, convertendo-a em hidroxilaminocitosina (**Figura 18.17 C**). Essa conversão aumenta a frequência de um tautômero raro que pareia com adenina em vez de guanina e leva a transições C • G → T • A. Como a hidroxilamina atua apenas sobre citosina, ela não gera transições T • A → C • G; dessa forma, ela não reverterá as mutações que produz. Resolva o Problema 25

Reações oxidativas. As espécies reativas do oxigênio (incluindo radicais superóxido, peróxido de hidrogênio e radicais

hidroxila) são produzidas durante o metabolismo aeróbico normal, assim como por radiação, ozônio, peróxidos e algumas drogas. Essas formas reativas de oxigênio danificam o DNA e induzem mutações ao provocarem mudanças químicas no DNA. Por exemplo, a oxidação converte guanina em 8-oxi-7,8-di-hidrodesoxiguanina (**Figura 18.18**), que frequentemente faz um pareamento incorreto com adenina em vez de citosina, provocando uma mutação de transversão G • C → T • A.

Agentes intercalantes. Proflavina, acridina laranja, brometo de etídio e dioxina são **agentes intercalantes** (**Figura 18.19 A**), que produzem mutações ao se intercalarem entre as bases adjacentes no DNA, distorcendo a estrutura tridimensional da hélice e provocando inserções e deleções de um nucleotídio na replicação (**Figura 18.19 B**). Essas inserções e deleções frequentemente produzem mutações no quadro de leitura, e os efeitos mutagênicos dos agentes intercalantes são graves. Como esses agentes geram adições e deleções, eles podem reverter os efeitos de suas próprias mutações.

Conceitos

As substâncias químicas podem produzir mutações por meio de vários mecanismos. Os análogos de base são incorporados ao DNA e frequentemente pareiam com a base errada. Os agentes alquilantes, substâncias químicas desaminadoras, hidroxilamina e espécies reativas de oxigênio mudam a estrutura das bases de DNA, alterando suas propriedades de pareamento. Os agentes intercalantes se instalam entre as bases e provocam inserções e deleções de base única na replicação.

Checagem dos conceitos 4

Qual característica indica que os análogos de base são mutagênicos?

a. Eles produzem mudanças na DNA polimerase que provocam seu mau funcionamento.
b. Eles distorcem a estrutura do DNA.
c. São semelhantes em estrutura às bases normais.
d. Provocam modificações químicas nas bases normais.

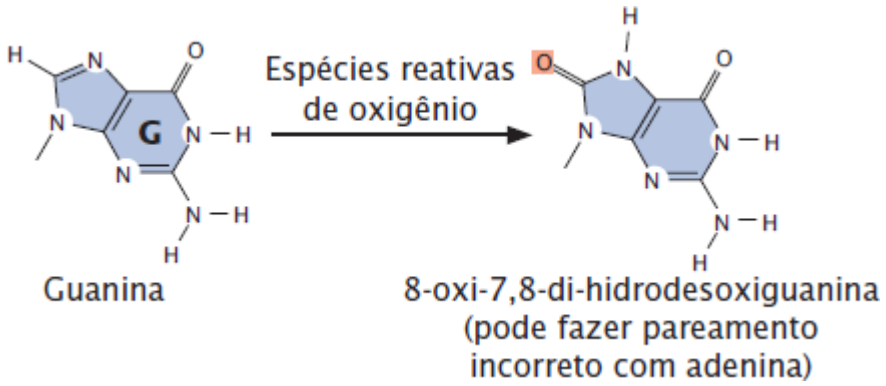


Figura 18.18 Espécies reativas de oxigênio convertem guanina em 8-oxi-7,8-di-hidrodesoxiguanina.

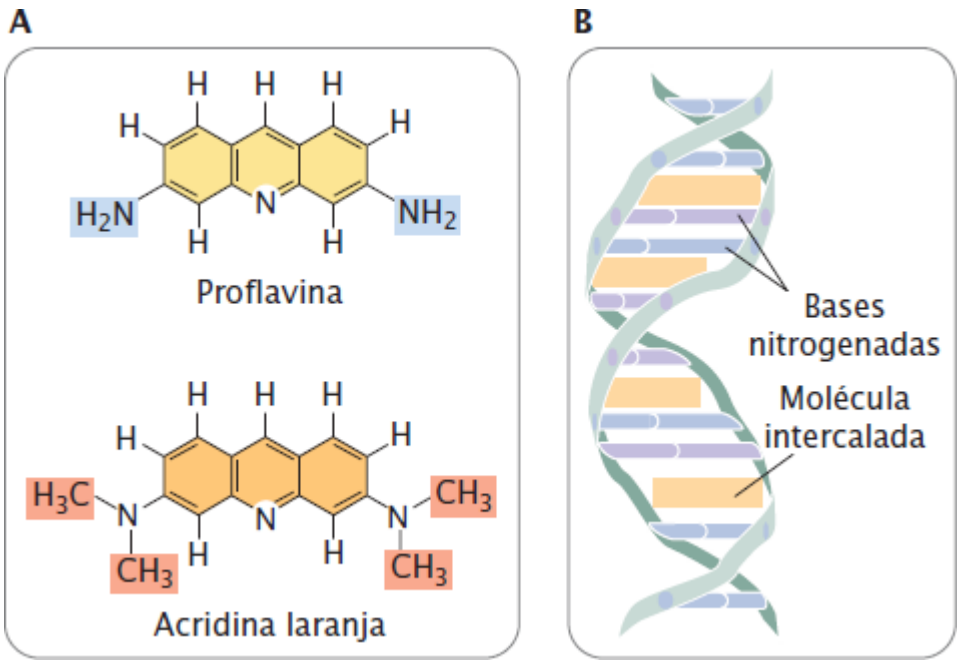


Figura 18.19 **Agentes intercalantes.** Proflavina e acridina laranja (**A**) se inserem entre as bases adjacentes no DNA, distorcendo a estrutura tridimensional da hélice (**B**).

Radiação

Em 1972, Hermann Muller demonstrou que as mutações nas moscas-da-fruta poderiam ser induzidas pelos raios X. Os resultados dos estudos subsequentes mostraram que os raios X aumentam em muito as taxas de mutação em todos os organismos. Por causa de sua energia elevada, os raios X, gama e cósmicos são capazes de penetrar nos tecidos e danificar o DNA. Essas formas de radiação, chamadas de radiação ionizante, expulsam os elétrons dos átomos onde eles se encontram, modificando as moléculas estáveis em radicais livres e íons reativos, que então alteram as estruturas de bases e quebram as ligações fosfodiéster no DNA. A radiação ionizante também provoca rupturas de fita dupla no DNA. As tentativas para reparar essas rupturas podem produzir mutações cromossômicas (discutidas no *Capítulo 8*).

A luz ultravioleta (UV) tem menos energia que a radiação ionizante e não ejeta os elétrons, mas mesmo assim, é extremamente mutagênica. As bases purina e pirimidina absorvem com facilidade a luz UV, resultando na formação de ligações químicas entre as moléculas pirimidina adjacentes na mesma fita de DNA e na criação de **dímeros de pirimidina** (**Figura 18.20 A**). Os dímeros de pirimidina compostos por duas bases timina (chamados de dímeros de timina) são mais frequentes, mas também podem surgir dímeros de citosina e timina-citosina. Os dímeros distorcem a configuração do DNA (**Figura 18.20 B**) e bloqueiam a replicação. A maioria dos dímeros de pirimidina é imediatamente reparada por mecanismos discutidos adiante neste capítulo, mas alguns escapam do reparo e inibem a replicação e a transcrição.

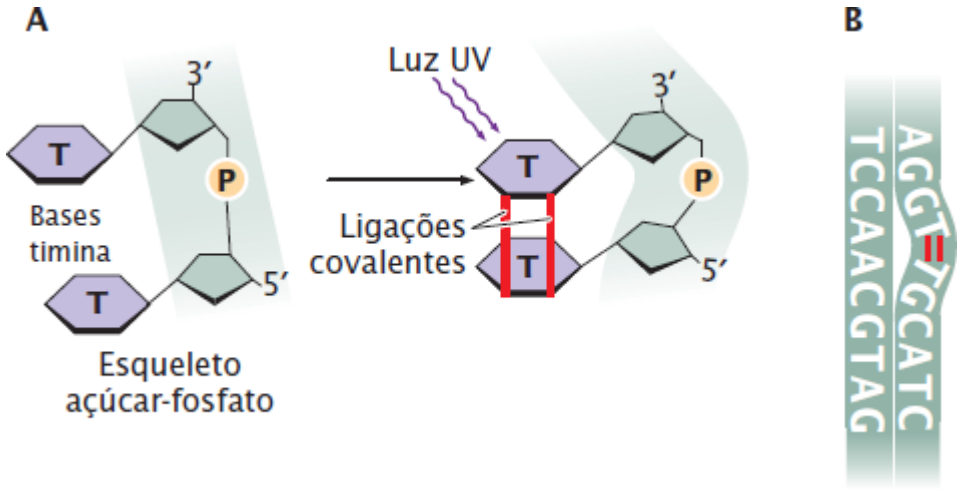


Figura 18.20 **Os dímeros de pirimidina são produzidos pela luz ultravioleta. A.** Formação do dímero da timina. **B.** DNA distorcido.

Quando os dímeros de pirimidina bloqueiam a replicação, a divisão celular é inibida e a célula em geral morre; por isso, a luz UV mata as bactérias e é um agente esterilizante eficiente. Para ocorrer uma mutação – um erro hereditário nas instruções genéticas –, o bloqueio da replicação deve ser superado. Às vezes as bactérias burlam os bloqueios de replicação produzidos

pelos dímeros de pirimidina e outros tipos de dano ao DNA por meio do **sistema SOS**. Esse sistema torna possível que os bloqueios de replicação sejam superados, mas no processo, produz vários erros e aumenta muito a taxa de mutação. De fato, o verdadeiro motivo para a replicação poder continuar se houver um bloqueio é que as enzimas no sistema SOS não obedecem rigorosamente às regras de pareamento de bases. A concessão é que a replicação pode continuar e a célula sobrevive, mas apenas à custa do sacrifício da precisão normal da síntese do DNA.

Conceitos

A radiação ionizante como os raios X e os raios gama danifica o DNA ao expulsar os elétrons dos átomos; esses elétrons, então, rompem as ligações fosfodiéster e alteram a estrutura das bases. A luz ultravioleta provoca mutações principalmente ao produzir dímeros de pirimidina que interrompem a replicação e a transcrição. O sistema SOS torna possível que as bactérias superem os bloqueios de replicação, mas introduz erros na replicação.

18.3 As mutações são objeto de intenso estudo pelos geneticistas

As mutações são estudadas pelos geneticistas por causa dos seus efeitos prejudiciais. Esses estudos incluíram o desenvolvimento de testes para determinar as propriedades mutagênicas dos compostos químicos e a investigação de populações humanas expostas de forma trágica a altos níveis de radiação.

Como detectar as mutações com o teste de Ames

As pessoas nas sociedades industriais estão circundadas por inúmeras substâncias químicas artificialmente produzidas: mais de 50.000 substâncias químicas diferentes estão no uso comercial e industrial atualmente, e de 500 a 1.000 substâncias químicas novas são introduzidas todo ano. Algumas dessas substâncias químicas são potenciais carcinógenos e muitos produtos naturais também são potencialmente carcinogênicos. Um método para testar o potencial das substâncias para provocar câncer é administrá-las a animais de laboratório (ratos ou camundongos) e comparar a incidência de câncer nos animais tratados com a incidência dos animais-controle. Infelizmente, esses testes são demorados e caros. Além disso, a capacidade de uma substância em provocar câncer em roedores nem sempre é indicativa de seu efeito nos humanos. Afinal de contas, não somos ratos!

Em 1974, Bruce Ames desenvolveu um teste simples para avaliar o potencial de substâncias químicas para provocar câncer. O **teste de Ames** baseia-se no princípio de que o câncer e as mutações são resultado de dano ao DNA, e os resultados dos experimentos demonstraram que 90% dos carcinógenos conhecidos também são mutágenos. Ames propôs que a mutagênese nas bactérias poderia servir como um indicador da carcinogênese nos humanos.

O teste de Ames usa diferentes cepas auxotróficas da bactéria *Salmonella typhimurium* com defeitos na camada de lipopolissacarídio, que normalmente protege as bactérias de substâncias químicas encontradas no ambiente. Além disso, o sistema de reparo do DNA nessas cepas foi inativado, aumentando sua suscetibilidade a mutágenos.

A versão mais recente do teste (chamado Ames II) usa várias cepas auxotróficas que detectam diferentes tipos de substituições de pares de base. Outras cepas detectam diferentes tipos de mutações no quadro de leitura. Cada cepa tem uma mutação *his*$^-$, que a torna incapaz de sintetizar o aminoácido histidina, e as bactérias são plaqueadas em um meio sem histidina (**Figura 18.21**). Apenas as bactérias que sofreram uma mutação reversa do gene histidina (*his*$^-$ → *his*$^+$) são capazes de sintetizar a histidina e crescem no meio, o que facilita a detecção dessas mutações. Diferentes diluições de uma substância química a ser testada são adicionadas a placas inoculadas com as bactérias, e o número de colônias bacterianas mutantes que surgem em cada placa é comparado com o número que aparece nas placas de controle sem substância química (ou seja, que surgem por meio da mutação espontânea). Qualquer substância química que aumente significativamente o número de colônias que aparece em uma placa tratada é mutagênica e provavelmente também carcinogênica.

Alguns compostos não são carcinógenos ativos, mas podem ser convertidos em compostos causadores de câncer no corpo. Para tornar o teste de Ames sensível para tais potenciais carcinógenos, um composto a ser testado é primeiro incubado no extrato de fígado de mamífero que contém enzimas metabólicas.

O teste de Ames é aplicado a milhares de substâncias químicas e produtos comerciais. Uma demonstração inicial da sua utilidade foi a descoberta, em 1975, de que muitas tinturas para cabelo vendidas nos EUA tinham compostos mutagênicos para as bactérias. Esses compostos foram removidos da maioria das tinturas.

Conceitos

O teste de Ames usa cepas *his*$^-$ de bactérias para testar a capacidade de substâncias químicas em produzir mutações *his*$^-$
$^+$

→ *his* . Como existe uma correlação entre a atividade mutagênica e o potencial carcinogênico, o teste de Ames é muito usado para classificar substâncias químicas quanto ao seu potencial para provocar câncer.

Experimento

Pergunta: Como as substâncias químicas podem ser rapidamente classificadas quanto a sua capacidade de provocar câncer?

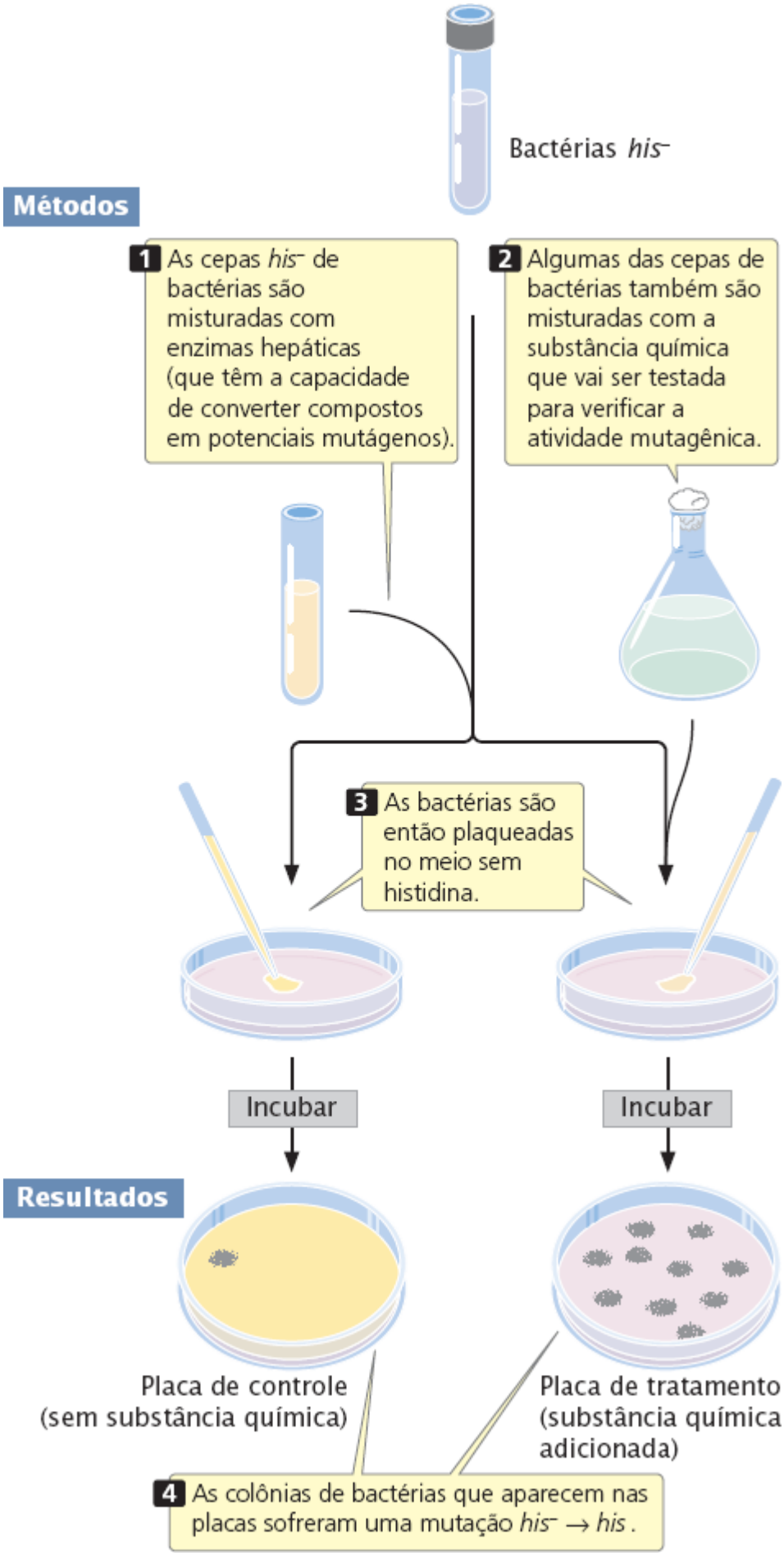


Figura 18.21 **O teste de Ames é usado para identificar mutágenos.**

Conclusão: Qualq uer substância química que aumente significativamente o número de colônias que aparece em uma placa de tratamento é mutagênica e provavelmente também carcinogênica.

Exposição de seres humanos à radiação

As pessoas são diariamente expostas a baixos níveis de radiação de fontes cósmicas, médicas e ambientais, mas existem eventos trágicos que produziram exposições em um nível muito mais elevado.

Em 6 de agosto de 1945, um avião norte-americano jogou uma bomba atômica na cidade de Hiroshima, no Japão. A explosão devastou a área da cidade que media 11,65 km^2, matou de 90.000 a 140.000 pessoas e afetou quase o mesmo número. Três dias depois, os EUA jogaram uma bomba atômica na cidade de Nagasaki, dessa vez destruindo uma área de 3,88 km^2 e matando entre 60.000 e 80.000 pessoas. Muita radiação foi liberada durante essas explosões e muitas pessoas foram expostas.

Após a guerra, um esforço unido Japão-EUA foi feito para estudar os efeitos biológicos da exposição à radiação nos sobreviventes das explosões atômicas e em seus filhos. As mutações somáticas foram examinadas ao estudar os efeitos da radiação e o câncer nos sobreviventes; foram avaliadas mutações germinativas ao observar os defeitos congênitos, anomalias cromossômicas e mutações gênicas nos filhos das pessoas que foram expostas à radiação.

O geneticista James Neel e seus colaboradores examinaram quase 19.000 filhos de pessoas que estiveram a 2.000 m do centro da explosão atômica em Hiroshima e Nagasaki, junto com um número semelhante de crianças cujos pais não foram expostos à radiação. As doses de radiação para os pais de uma criança foram estimadas com base na cuidadosa avaliação da localização, postura e posição dos pais no momento da explosão. Foi coletada uma amostra de sangue de cada criança e foi usada eletroforese em gel para investigar as substituições de aminoácidos em 28 proteínas. Quando eram detectadas variantes raras, as amostras de sangue dos pais da criança também eram analisadas para estabelecer se a variante era herdada ou se era uma nova mutação.

De um total de 289.868 genes examinados por Neel e seus colaboradores, foi encontrada apenas uma mutação nos filhos de pessoas expostas; não foram encontradas mutações no grupo-controle. A partir desses achados, estimou-se uma taxa de mutação de $3{,}4 \times 10^{-6}$ para crianças cujos pais foram expostos à explosão, que está dentro da faixa de taxas de mutação espontânea observadas em outros eucariotos. Neel e seus colaboradores também examinaram as frequências das mutações cromossômicas, as razões sexuais dos filhos de pessoas expostas e as frequências de aneuploidia cromossômica. Não foram encontradas evidências em nenhum desses ensaios de aumento das mutações nos filhos das pessoas expostas à radiação das explosões atômicas, sugerindo que as mutações germinativas não estavam elevadas.

Estudos com animais mostraram claramente que a radiação provoca mutações germinativas; então, por que não havia aumento aparente nessas mutações nos habitantes de Hiroshima e Nagasaki? Os pais expostos exibiam maior incidência de leucemia e outros tipos de câncer; então, estava claro que houve indução das mutações somáticas. A resposta para a pergunta é desconhecida, mas a falta de mutações germinativas pode ocorrer pelo fato de que as pessoas que receberam as maiores doses de radiação morreram logo após a explosão. Um esclarecimento adicional sobre os efeitos genéticos da radiação veio a partir de estudos de pessoas expostas à radiação no acidente nuclear de Chernobyl, em 1986, e outros acidentes nucleares, assim como a exposição à radiação usada na medicina e na indústria.

18.4 Os elementos t de transposição provocam mutações

Os **elementos de transposição** – sequências que podem se deslocar no genoma – também podem ser responsáveis pelas mutações. Esses elementos móveis de DNA têm vários nomes, incluindo transpósons, elementos genéticos de transposição, genes móveis, elementos controladores e genes *jumping*. Eles são encontrados nos genomas de todos os organismos e são abundantes em muitos: por exemplo, compõem pelo menos 45% do DNA humano. A maioria dos elementos de transposição é capaz de se inserir em muitos locais diferentes, dependendo de mecanismos diferentes da recombinação homóloga. Por meio do seu movimento (transposição), esses elementos provocam mutações, ao se inserir em outro gene e rompendo-o ou ao promover rearranjos do DNA como deleções, duplicações e inversões cromossômicas (ver *Capítulo 8*).

Características gerais dos elementos de transposição

Existem muitos tipos diferentes de elementos de transposição: alguns têm estruturas simples, incluindo apenas as sequências necessárias para sua própria transposição, enquanto outros apresentam estruturas complexas e codificam várias funções não diretamente relacionadas com a transposição. Apesar dessa variação, muitos elementos de transposição têm algumas características em comum.

As **repetições diretas flanqueadoras** curtas de 3 a 12 pb são encontradas em ambos lados da maioria dos elementos de transposição. As sequências dessas repetições variam, mas o comprimento é constante para cada tipo de elemento de transposição. Essas repetições não são uma parte de um elemento de transposição e não viajam com ele. Em vez disso, elas são geradas no processo de transposição, no ponto de inserção. Durante a transposição, as repetições flanqueadoras são criadas quando são feitos cortes escalonados no DNA-alvo, como apresentado na **Figura 18.22**. Os cortes escalonados deixam

segmentos curtos de fita única de DNA em ambos os lados do elemento de transposição. A replicação do DNA de fita única, então, cria as repetições diretas flanqueadoras.

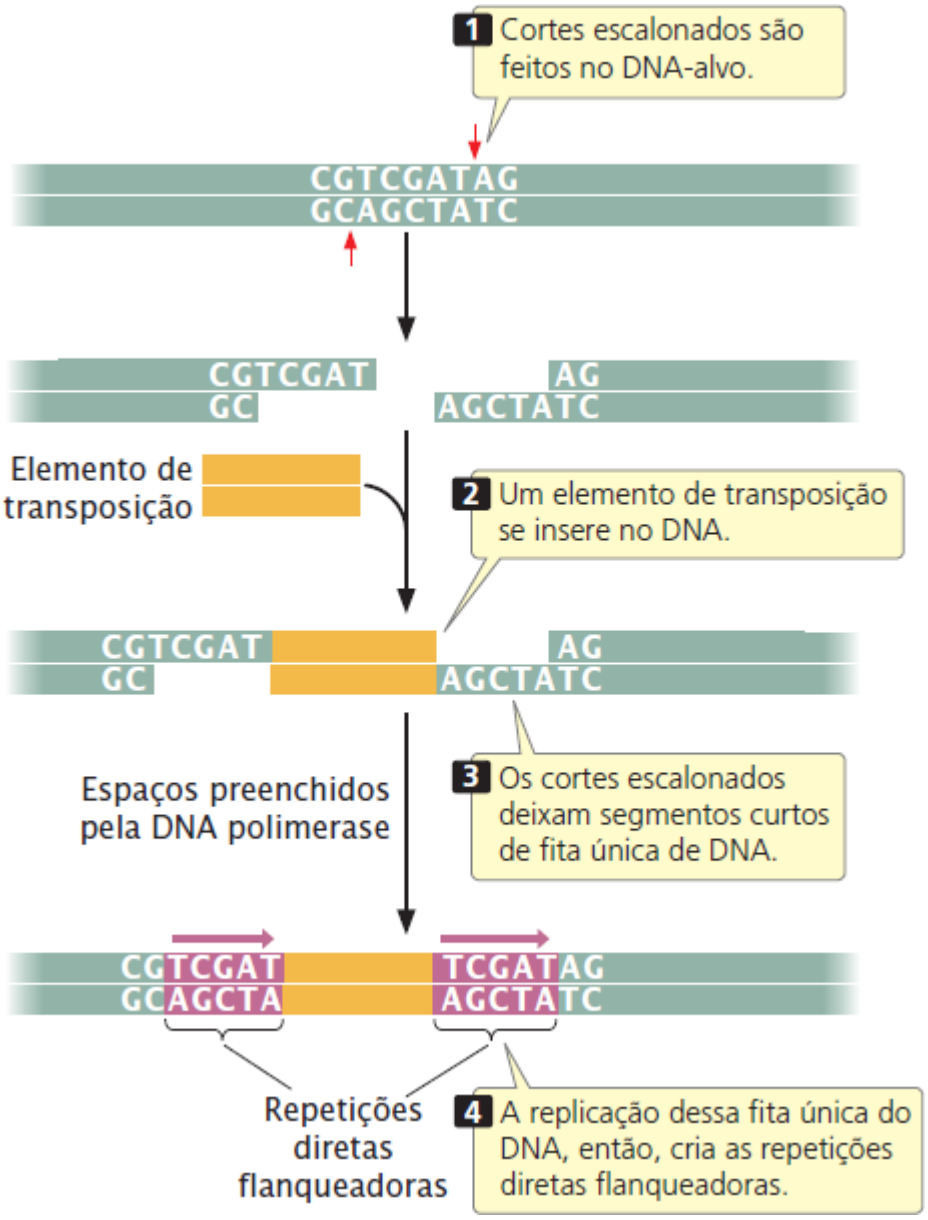


Figura 18.22 As repetições diretas flanqueadoras são geradas quando um elemento de transposição se insere no DNA.

Nas extremidades de muitos, mas não de todos, elementos de transposição são encontradas **repetições invertidas terminais** – sequências de 9 a 40 pb de comprimento que são complementos invertidos um do outro. Por exemplo, as sequências a seguir são repetições invertidas:

5′-ACAGTTCAG ... CTGAACTGT-3′
3′-TGTCAAGTC ... GACTTGACA-5′

Na mesma fita, as duas sequências não são inversões simples, como seu nome poderia implicar; pelo contrário, elas são invertidas e complementares. (Observe que a sequência da esquerda para a direita na fita superior é a mesma que a sequência da direita para esquerda da fita inferior.) As repetições invertidas terminais são identificadas por enzimas que catalisam a transposição e são necessárias para que ocorra a transposição. A **Figura 18.23** resume as características gerais dos elementos de transposição. Resolva o Problema 33

Conceitos

Os elementos de transposição são sequências móveis de DNA que provocam mutações. Existem muitos tipos diferentes de elementos de transposição; a maioria gera repetições diretas flanqueadoras nos sítios-alvo à medida que se insere. Muitos elementos de transposição também têm repetições invertidas terminais

curtas.

Checagem dos conceitos 5

Como as repetições diretas flanqueadoras são criadas na transposição?

Transposição

Como já foi citado, a **transposição** é o movimento de um elemento de transposição de um local para outro. São usados vários mecanismos diferentes para transposição nas células procarióticas e eucarióticas. Contudo, todos os tipos de transposição têm algumas características em comum: (1) são feitas quebras escalonadas no DNA-alvo (ver Figura 18.22); (2) o elemento de transposição é unido a extremidades de fita única do DNA-alvo e (3) o DNA é replicado nas lacunas na fita única. Uma enzima **transposase**, codificada pelo elemento de transposição, é usada para criar as quebras escalonadas no DNA e para integrar o elemento de transposição em um novo sítio, e ambos fazem com que o elemento de transposição se desloque.

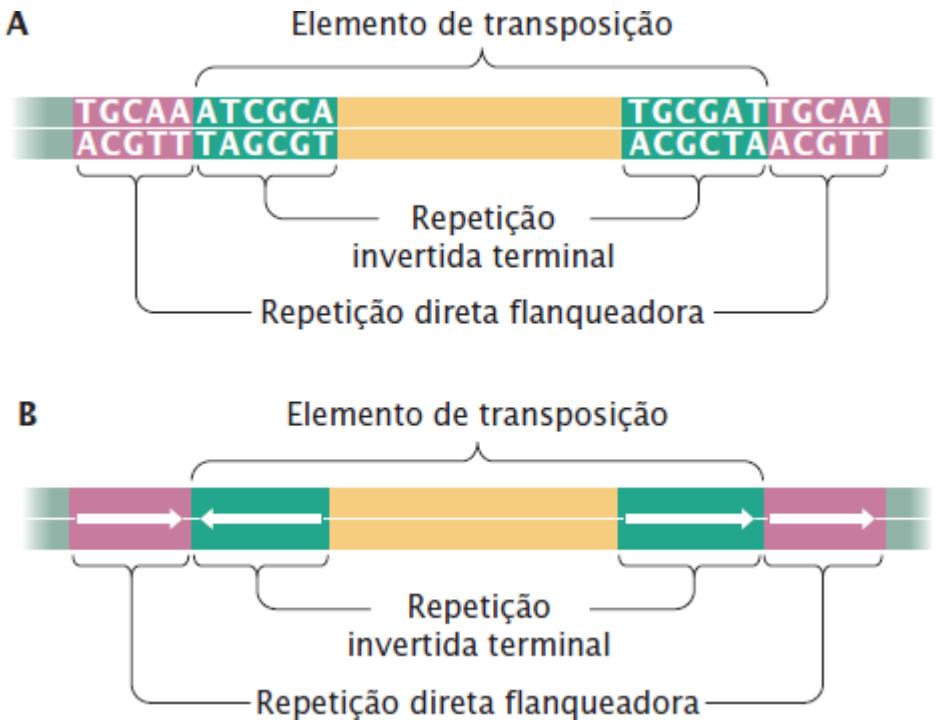


Figura 18.23 Muitos elementos de transposição têm características comuns. A. A maioria dos elementos de transposição gera repetições diretas flanqueadoras em cada lado do ponto de inserção no DNA-alvo. Muitos elementos de transposição também têm repetições terminais invertidas. **B.** Estas representações de repetições diretas e indiretas são usadas nas ilustrações deste capítulo.

Alguns elementos de transposição se transpõem ao DNA e são chamados de **transpósons de DNA** (também conhecidos como elementos de transposição Classe II). Outros elementos de transposição se transpõem por meio de um RNA intermediário. Nesse caso, o RNA é transcrito a partir do elemento de transposição (DNA) e então copiado de volta a DNA por uma enzima especial chamada transcriptase reversa. Os elementos que transpõem por meio de um RNA intermediário são chamados de **retrotranspósons** (também conhecidos como transpósons Classe I). A maioria dos elementos de transposição encontrada nas bactérias são transpósons de DNA. Ambos transpósons e retrotranspósons são encontrados nos eucariotos, embora os retrotranspósons sejam mais comuns.

Entre os transpósons de DNA, a transposição pode ser replicativa ou não replicativa. Na **transposição replicativa** (também chamada de transposição copia e cola), uma nova cópia do elemento de transposição é introduzida em um novo sítio enquanto a cópia antiga permanece atrás do sítio original, e então o número de cópias do elemento de transposição aumenta como resultado da transposição. Na **transposição não replicativa** (transposição corta e cola), o elemento de transposição é cortado do sítio antigo e se insere em um novo sítio sem nenhum aumento no número de suas cópias. A transposição não replicativa requer a replicação de apenas alguns nucleotídios que compõem as repetições diretas. Os retrotranspósons usam apenas a transposição replicativa.

Controle da transposição. Muitos organismos limitam a transposição ao metilar o DNA em regiões onde os transpósons são comuns. A metilação do DNA em geral suprime a transcrição (ver *Capítulo 17*), evitando a produção da enzima transposase necessária para a transposição. As alterações na estrutura da cromatina também são usadas para evitar a transcrição dos

transpósons. Em outros casos, a tradução do mRNA da transposase é controlada. Alguns animais usam pequenas moléculas de RNA chamadas de RNAs que interagem com Piwi (piRNAs, ver *Capítulo 14*), para silenciar os transpósons; os piRNAs se combinam com as proteínas Piwi e inibem a expressão das sequências de transpósons.

Transposição em humanos. Cerca de 45% do genoma humano está restrito a sequências relacionadas com elementos de transposição, a sua maioria retrotranspósons (pesquisas sugerem que quase dois terços do genoma humano é composto de elementos de transposição). Os pesquisadores acreditavam que a maioria desses elementos de transposição fosse inativa e que ocorresse pouca transposição atualmente, embora esteja claro que ela tenha ocorrido em uma grande extensão durante a evolução no passado. Mais recentemente, entretanto, os pesquisadores começaram a mapear cópias de transpósons pelo genoma e descobriram que o número e a localização dos transpósons variam entre as pessoas. Isso sugere que as transposições recentes são mais comuns do que se acreditava. Por exemplo, estima-se que o transpóson L1 sofra um evento de transposição a cada 100 nascimentos.

Pesquisas também demonstraram que algumas células de câncer têm níveis elevados de transposição, provavelmente por causa de os padrões de metilação do DNA – que normalmente inibem a transposição – estarem interrompidos nessas células.

Conceitos

A transposição pode ocorrer por meio de um DNA ou RNA intermediário. Na transposição replicativa, uma nova cópia de elemento de transposição se insere em um novo local e a cópia antiga permanece para trás; na transposição não replicativa, a cópia antiga é retirada do sítio antigo e se desloca para um novo sítio. A transposição por meio de um RNA intermediário requer a transcrição reversa para se integrar no sítio-alvo. Muitas células regulam a transposição por meio de vários mecanismos.

Efeitos mutagênicos da transposição

Como os elementos de transposição podem se inserir em outros genes e atrapalhar sua função, a transposição é geneticamente mutagênica. Na verdade, mais da metade de todas as mutações de ocorrência espontânea na *Drosophila* é resultado da inserção de um elemento de transposição em um gene funcional ou próximo dele.

Foram rastreados vários casos de doenças genéticas humanas pela inserção de um elemento genético de transposição em um gene vital. Por exemplo, a inserção do elemento de transposição L1 no gene para o fator VIII da coagulação sanguínea provoca a hemofilia. Embora a maioria das mutações resultantes da transposição seja prejudicial, a transposição pode ativar um gene ou mudar o fenótipo da célula de forma vantajosa em alguns casos. Por exemplo, os elementos de transposição bacterianos carreiam genes que codificam a resistência aos antibióticos, e vários elementos de transposição criaram mutações que conferem resistência aos inseticidas nos insetos.

Um exemplo drástico do efeito mutagênico dos elementos de transposição é a cor das uvas, que existe nas variedades preta, vermelha e branca (**Figura 18.24**). As uvas pretas e vermelhas são o resultado da produção dos pigmentos vermelhos (antocianina) na casca, que não existem nas uvas brancas. As uvas brancas são resultado de uma mutação nas uvas pretas que desliga a produção do pigmento antocianina. Essa mutação é composta pela inserção de um retrotranspóson de 10.422 pb chamado *Gret1* próximo de um gene que promove a produção das antocianinas. A retrotransposição *Gret1* aparentemente rompe as sequências que regulam o gene, efetivamente desligando a produção de pigmento e produzindo uma uva branca sem antocianinas. As uvas vermelhas resultam de uma segunda mutação que ocorre nas uvas brancas (ver Figura 18.24). Essa mutação (provavelmente resultante de uma recombinação defeituosa) removeu quase, mas não todos, os retrotranspósons, ligando a produção de pigmento de volta, mas não de forma tão intensa como as uvas pretas originais.

Como a transposição leva à troca de sequências de DNA e recombinação, ela leva a rearranjos do DNA. A recombinação homóloga entre múltiplas cópias de transpósons pode levar a duplicações, deleções e inversões, como apresentado na **Figura 18.25**. A mutação *Bar* na *Drosophila* (ver Figura 8.5) é uma duplicação em *tandem* que se acreditava ter surgido da recombinação homóloga entre duas cópias de um elemento de transposição encontrado em diferentes locais no cromossomo X. Da mesma forma, a recombinação entre as cópias do elemento de transposição *Rider* provocou uma duplicação que resultou na fruta alongada nos tomates.

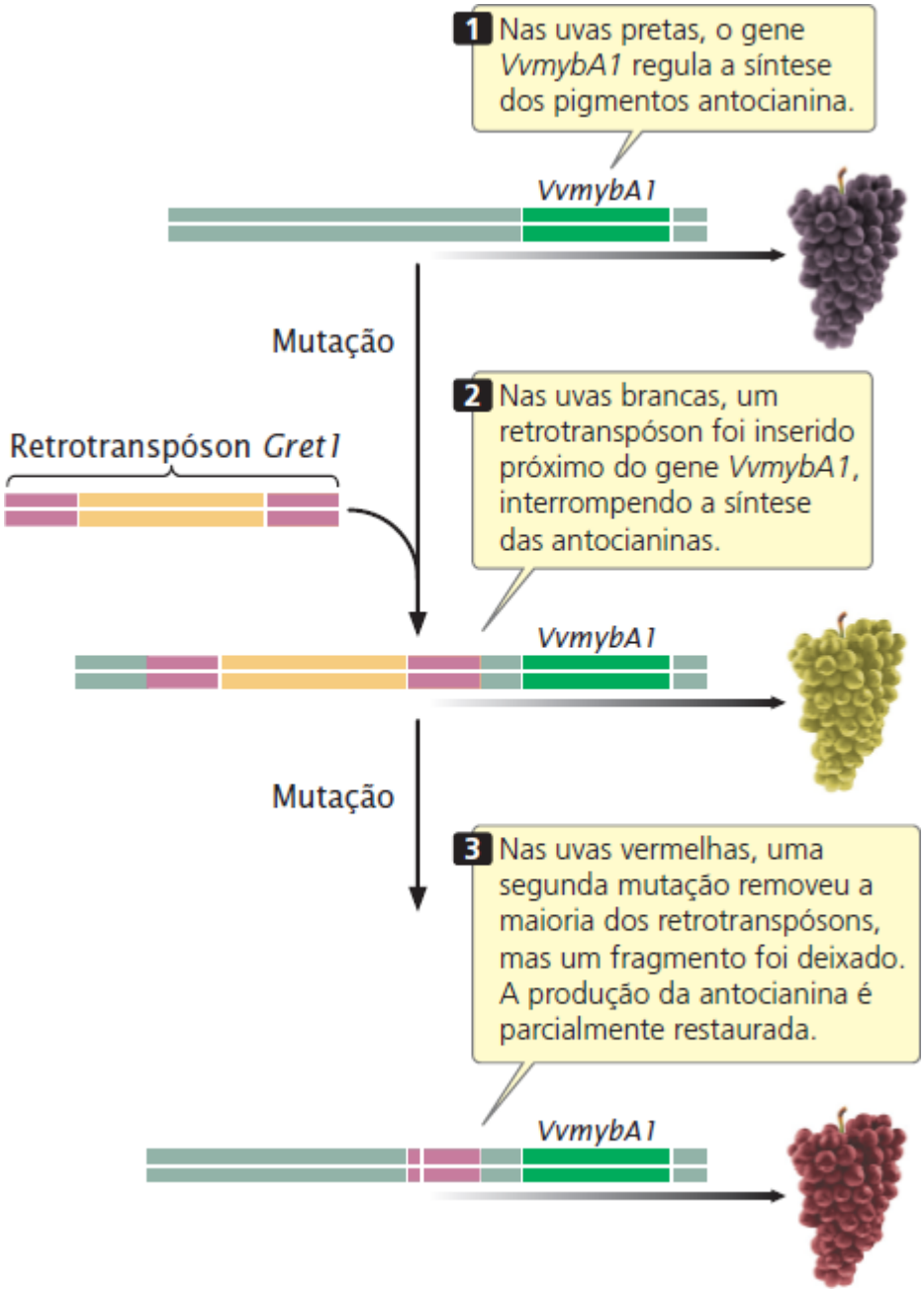
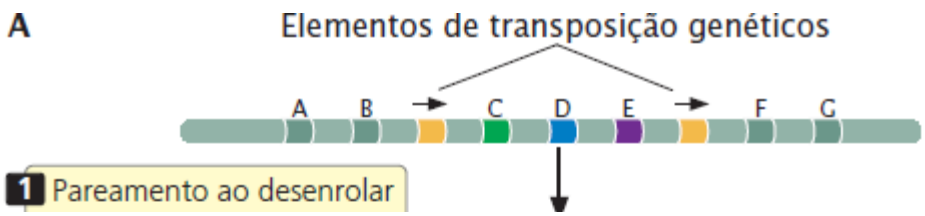
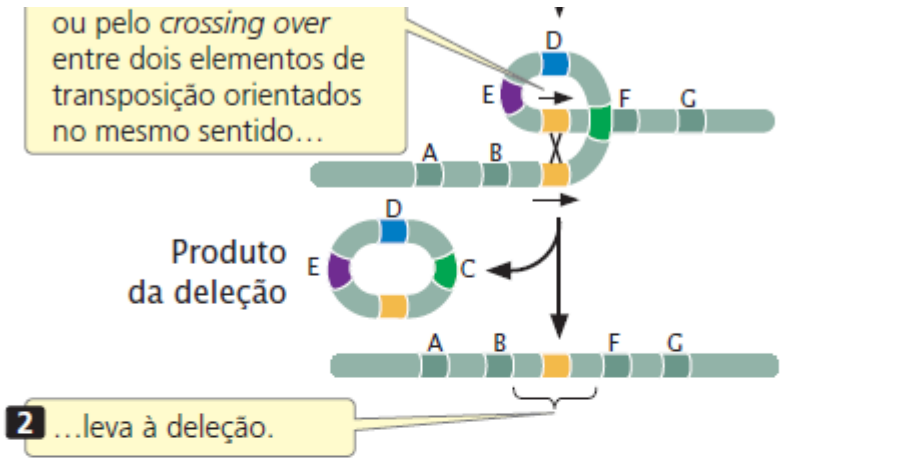


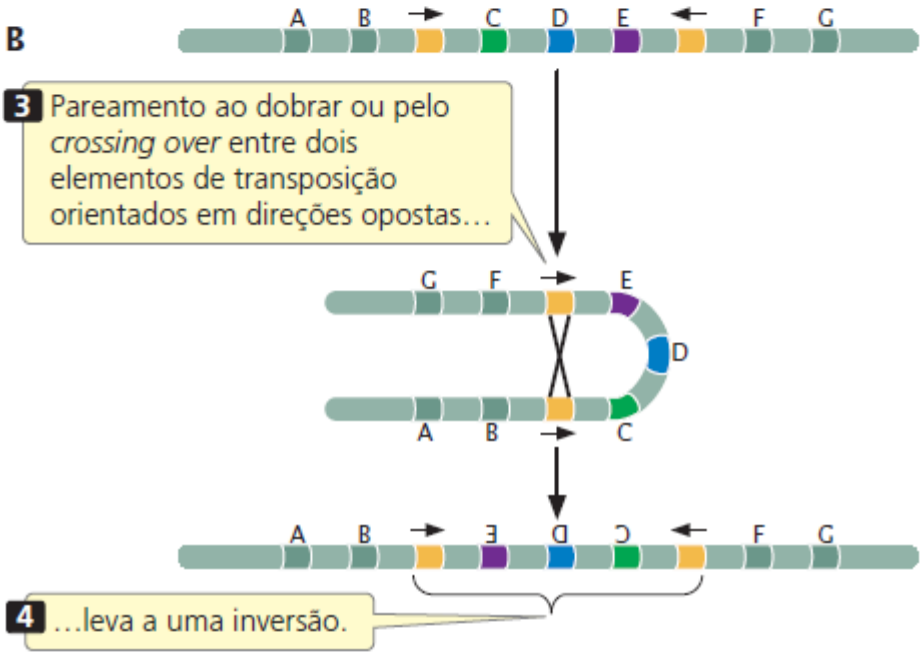
Figura 18.24 As cores vermelha e branca nas uvas são resultado da inserção e deleção de um retrotranspóson.

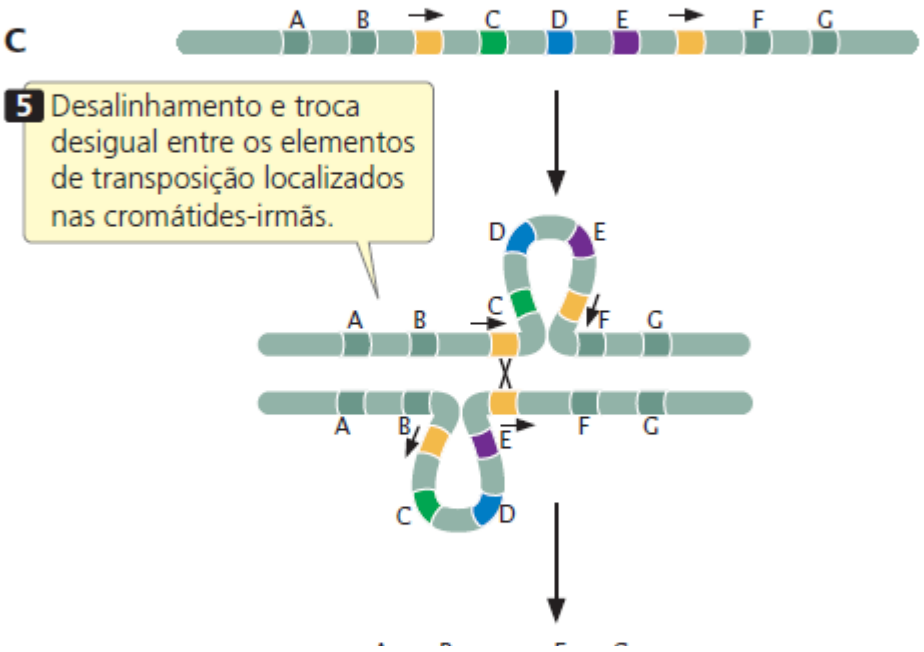
Os rearranjos de DNA também podem ser causados pela retirada dos elementos de transposição em uma transposição corta e cola. Se o DNA rompido não é adequadamente reparado, pode ser gerado um rearranjo cromossômico.

Como a maioria dos elementos de transposição se insere de forma aleatória nas sequências de DNA, eles são ferramentas úteis aos pesquisadores para induzir mutações por todo o genoma, possibilitando determinar as funções dos genes, estudar os fenômenos genéticos e mapear os genes. Além disso, como o elemento de transposição a ser usado tem uma sequência conhecida, ele pode servir como um "adesivo" para localizar o gene no qual a mutação ocorreu. Por exemplo, os pesquisadores desenvolveram um elemento de transposição chamado *Bela Adormecida* para induzir mutações em camundongos e usá-lo para pesquisar os genes que provocam câncer. O *Bela Adormecida* foi introduzido em uma cepa de camundongos que produzem a transposase necessária para transposição, e o elemento de transposição foi inserido aleatoriamente em diferentes locais no genoma. Em alguns casos, ele se insere em um gene que protege contra o câncer e destrói sua função. Ao procurar a localização da sequência da *Bela Adormecida* no DNA das células tumorais que se desenvolvem, os geneticistas identificaram vários genes que protegem contra o câncer.




ou pelo *crossing over* entre dois elementos de transposição orientados no mesmo sentido…
D
E
F
G
A
B
Produto da deleção
D
E
C
A
B
F
G
2 …leva à deleção.


B
A
B
C
D
E
F
G
3 Pareamento ao dobrar ou pelo *crossing over* entre dois elementos de transposição orientados em direções opostas…
G
F
E
D
A
B
C
4 …leva a uma inversão.


C
A
B
C
D
E
F
G
5 Desalinhamento e troca desigual entre os elementos de transposição localizados nas cromátides-irmãs.
D
E
A
B
C
F
G
A
B
E
F
G
C
D

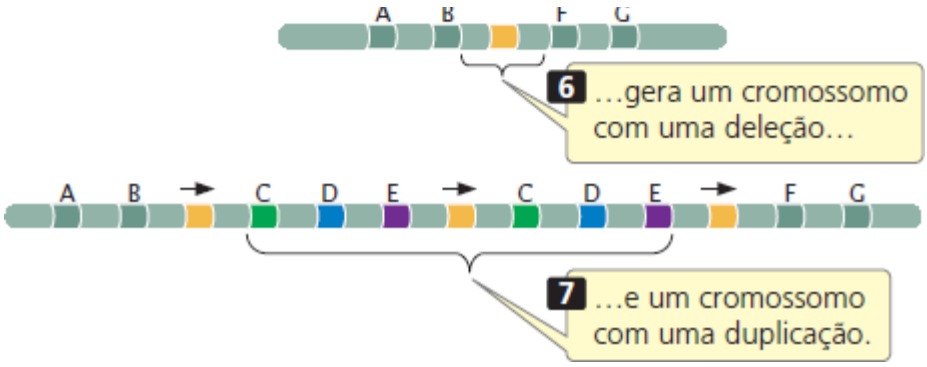


Figura 18.25 **Muitos rearranjos cromossômicos são gerados por transposição.**

As bactérias e os organismos eucarióticos têm vários tipos diferentes de elementos de transposição, e suas estruturas variam muito. Nas duas seções a seguir, vamos analisar a estrutura e os tipos de elementos de transposição encontrados nas bactérias e eucariotos.

Conceitos

Os elementos de transposição frequentemente provocam mutações e rearranjos de DNA.

Checagem dos conceitos 6

Explique em poucas palavras como a transposição provoca mutações e rearranjos cromossômicos.

Elementos de transposição nas bactérias

Os transpósons de DNA encontrados nas bactérias (não existem retrotranspósons nas bactérias) são compostos por dois grupos principais: (1) elementos de transposição simples, chamados de sequências de inserção, que carreiam apenas a informação necessária para o movimento e (2) elementos de transposição mais complexos, chamados de transpósons compostos, que têm sequências de DNA não relacionadas diretamente com a transposição.

Sequências de inserção. O tipo mais simples de elemento de transposição nos cromossomos bacterianos e plasmídios é uma **sequência de inserção** (IS; do inglês, *insertion sequence*). Esse tipo de elemento carreia apenas a informação genética necessária para seu movimento. As sequências de inserção são componentes comuns das bactérias; elas também podem infectar plasmídios e vírus e, dessa forma, são transmitidas de uma célula para outra. Os geneticistas designaram cada tipo de sequência de inserção com *IS* seguida com um número de identificação. Por exemplo, *IS1* é uma sequência de inserção comum encontrada na *E. coli*.

Já foram encontradas várias sequências de inserção diferentes nas bactérias. Elas têm tipicamente 800 a 2.000 pb de comprimento e as duas características dos elementos de transposição: repetições invertidas terminais e a geração de repetições diretas flanqueadoras no sítio de inserção. A maioria das sequências de inserção tem um ou dois genes que codificam transposase. A *IS1*, uma sequência de inserção típica, é apresentada na **Figura 18.26**. Resolva o Problema 36

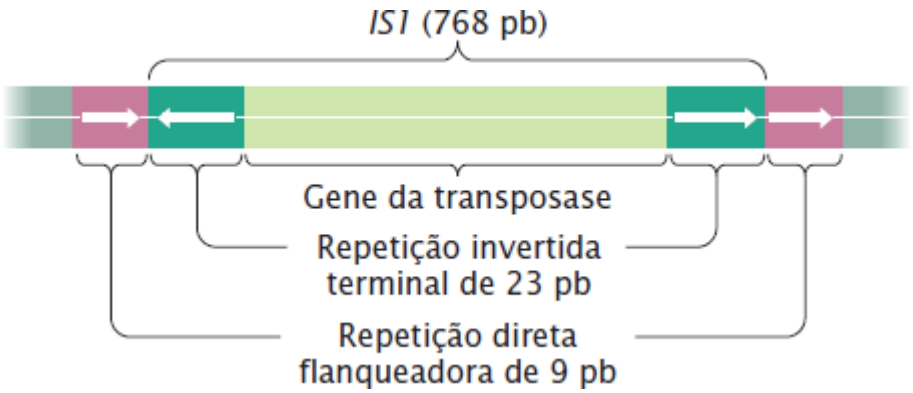


Figura 18.26 **As sequências de inserção são elementos de transposição simples encontrados nas bactérias.**

Transpósons compostos. Qualquer segmento de DNA que é flanqueado por duas cópias de uma sequência de inserção pode, por si só, se transpor – um fenômeno chamado de **transpóson composto**. Cada transpóson composto é designado por uma

abreviação *Tn*, seguido por um número. O transpóson composto *Tn10*, por exemplo, é composto por cerca de 9.300 pb que carreiam um gene para resistência à tetraciclina entre suas sequências de inserção *IS10* (**Figura 18.27**). As sequências de inserção têm repetições invertidas terminais, então o transpóson composto também termina em repetições invertidas. Os transpósons compostos também geram repetições diretas flanqueadoras em seus sítios de inserção (ver Figura 18.27). As sequências de inserção nas extremidades de um transpóson composto também podem estar na mesma orientação ou podem estar invertidas em relação a outra (como no *Tn10*).

As sequências de inserção nas extremidades de um transpóson composto são responsáveis pela transposição. O DNA entre as sequências de inserção não é necessário para o movimento e pode carrear informação adicional (como resistência a antibiótico). Aparentemente, os transpósons compostos se desenvolvem quando uma sequência de inserção se transpõe para um local próximo a outro do mesmo tipo. A transposase produzida por uma das sequências de inserção catalisa a transposição de ambas sequências de inserção, fazendo com que elas se movam juntas e carreiem o DNA que está entre elas. Em alguns transpósons compostos (como *Tn10*), uma das sequências de inserção pode ser defeituosa, então seu movimento depende da transposase produzida por outra.

Transpósons não compostos. Alguns elementos de transposição nas bactérias não têm sequências de inserção e são chamados de transpósons não compostos. Os transpósons não compostos têm um gene para transposase e têm repetições invertidas nas suas extremidades. Por exemplo, o transpóson não composto *Tn3* carreia os genes para a transposase e resolvase (uma enzima que atua na recombinação), mais um gene que codifica a enzima betalactamase, que fornece resistência ao antibiótico ampicilina.

Alguns genomas de bacteriófagos se reproduzem por transposição e usam a transposição para se inserir em um cromossomo bacteriano no seu ciclo lisogênico; o bacteriófago de transposição mais bem estudado é o Mu (**Figura 18.28**). Embora o Mu não tenha repetições invertidas terminais, ele gera repetições diretas flanqueadoras curtas (5 pb) quando se insere aleatoriamente no DNA. O Mu se replica por transposição e provoca mutações no sítio da inserção, propriedades características dos elementos de transposição.

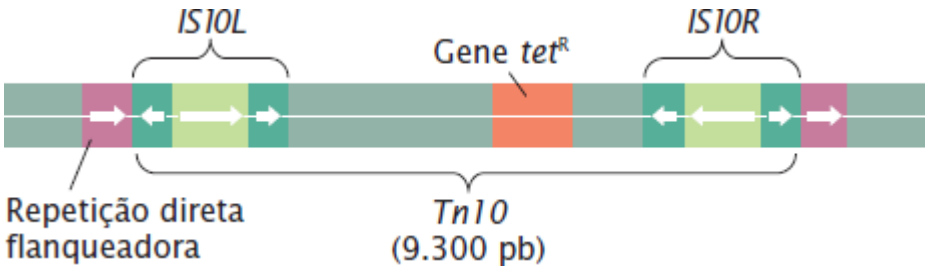


Figura 18.27 ***Tn10*** **é um transpóson composto encontrado nas bactérias.**

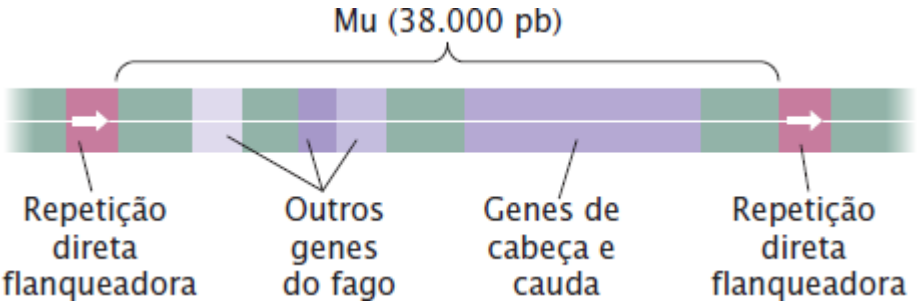


Figura 18.28 **Mu é um bacteriófago de transposição.**

Conceitos

As sequências de inserção são elementos de transposição procarióticos que carreiam apenas a informação necessária para a transposição. Um transpóson composto é composto por duas sequências de inserção mais um DNA interveniente. Os transpósons não compostos nas bactérias não têm sequências de inserção, mas têm repetições invertidas terminais e carreiam informações não relacionadas com a transposição. Todos esses elementos de transposição geram repetições diretas flanqueadoras em seus pontos de inserção.

Checagem dos conceitos 7

Qual tipo de elemento de transposição tem repetições invertidas terminais?

a. Sequência de inserção.
b. Transpósons compostos.
c. Transpóson não composto *Tn3*.
d. Todas as opções acima.

Elementos de transposição nos eucariotos

Os elementos de transposição eucarióticos podem ser divididos em dois grupos. Um grupo é estruturalmente semelhante aos elementos de transposição nas bactérias, termina em repetições invertidas curtas e se transpõe como o DNA: os exemplos incluem os elementos *P* na *Drosophila* e os elementos *Ac* e *Ds* no milho. O outro grupo inclui os retrotranspósons; eles usam RNAs intermediários e muitos são semelhantes em estrutura e movimento aos retrovírus (ver *Capítulo 9*). Com base na sua estrutura, função e sequências genômicas, alguns retrotranspósons têm uma clara relação evolutiva com os retrovírus. Embora seu mecanismo de movimento seja fundamentalmente diferente dos outros elementos de transposição, os retrotranspósons também geram repetições diretas no ponto de inserção. Os retrotranspósons incluem os elementos *Ty* na levedura, os elementos *copia* na *Drosophila* e as sequências *Alu* nos humanos.

Elementos Ac e Ds no milho. Os elementos de transposição foram identificados no milho pela primeira vez há mais de 50 anos por Barbara McClintock (**Figura 18.29**). McClintock passou boa parte de sua carreira estudando suas propriedades, e seu trabalho está entre as descobertas fundamentais da genética. Os resultados de seu estudo, entretanto, foram mal interpretados e ignorados por muitos anos. Não até as técnicas moleculares serem desenvolvidas, entre os anos 1960 e 1970, e a importância dos elementos de transposição serem aceitos. A importância das descobertas iniciais de McClintock foi finalmente reconhecida em 1983, quando ela recebeu o prêmio Nobel em Fisiologia ou Medicina.

A descoberta de McClintock sobre os elementos de transposição tem sua origem no trabalho inicial de Rollins A. Emerson sobre os genes do milho que geravam grãos variados (multicoloridos). A maioria dos grãos do milho é completamente pigmentada ou incolor (amarelo), mas Emerson observou que alguns grãos amarelos tinham pontos ou linhas de outra cor. Ele propôs que esses grãos eram o resultado de uma mutação instável: uma mutação no gene do tipo selvagem para o pigmento produzia um grão incolor, mas em algumas células, a mutação reverteu de volta para o tipo selvagem, gerando uma mancha de pigmento. Entretanto, Emerson não sabia por que essas mutações eram instáveis.

Figura 18.29 Barbara McClintock foi a primeira pesquisadora a descobrir os elementos de transposição. (Topham/The Image Works.)

McClintock descobriu que a causa da mutação instável era um gene que se deslocava. Ela observou que a ruptura do cromossomo no milho ocorria em um gene que ela chamou *Dissociação* (*Ds*), mas apenas se houvesse outro gene, o *Ativador*

(*Ac*). Em alguns casos, esses genes se deslocavam juntos para um local diferente no cromossomo. McClintock chamou esses genes móveis de elementos controladores, porque eles controlavam a expressão de outros genes.

Os elementos *Ac* e *Ds* no milho foram examinados minuciosamente desde que foi reconhecida a importância do trabalho de McClintock. Existem transpósons de DNA com repetições invertidas terminais e repetições diretas flanqueadoras nos pontos de inserção (**Figura 18.30 A**). Cada elemento *Ac* tem um único gene que codifica uma enzima transposase. Assim, os elementos *Ac* são *autônomos* – capazes de se transpor. Os elementos *Ds* são elementos *Ac* com uma ou mais deleções que inativaram o gene transposase (**Figura 18.30 B**). Incapazes de se transpor sobre si mesmos (*não autônomo*), os elementos *Ds* podem se transpor se forem encontrados elementos *Ac* porque eles ainda têm repetições invertidas terminais reconhecidas pela *Ac* transposase.

Cada grão em uma espiga de milho é um descendente individual, originado a partir de um óvulo fertilizado por um grão de pólen. O padrão de pigmento da semente é determinado por vários *loci*. Um alelo que codifica o pigmento em um desses *loci* pode ser designado como *C*, e um alelo no mesmo *locus* que não confere pigmento pode ser designado como *c*. Um grão com o genótipo *cc* será incolor – ou seja, amarelo ou branco (**Figura 18.31 A**); um grão com genótipo *CC* ou *Cc* produzirá pigmento e será púrpura (**Figura 18.31 B**).

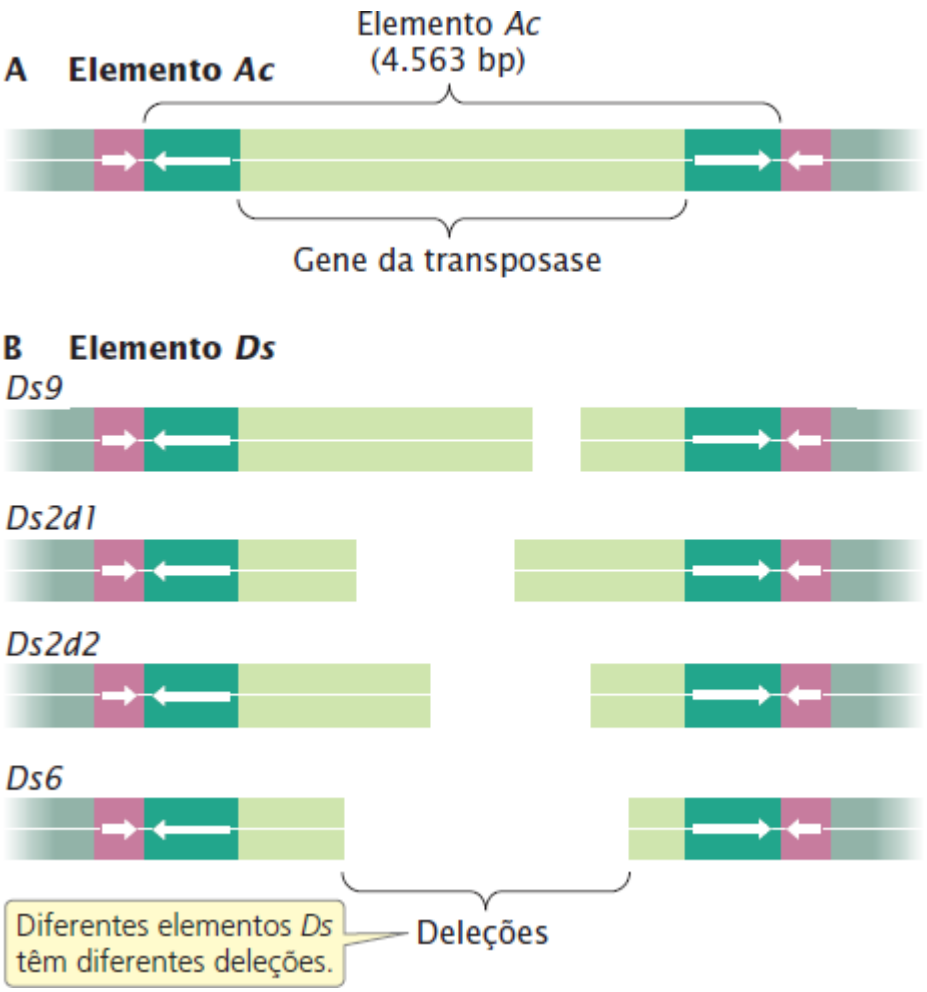


Figura 18.30 Os elementos *Ac* e *Ds* são elementos de transposição no milho.

Um elemento *Ds*, transposto sob a influência de um elemento *Ac* próximo, pode se inserir no alelo *C*, destruindo sua capacidade de produzir pigmento (**Figura 18.31 C**). Um alelo inativado por um elemento de transposição é chamado de "t" subscrito; então, nesse caso, o alelo seria C_t.

Se um grão for inicialmente heterozigoto com genótipo *Cc*, então após a transposição de *Ds* no alelo *C*, o genótipo dessa semente é C_tc. Esse grão será incolor (branco ou amarelo), porque nem o alelo C_t nem o alelo *c* conferem pigmento. À medida que o embrião de uma célula do milho se desenvolve e se divide por mitose, podem ocorrer transposições adicionais em algumas células. Em qualquer célula na qual o elemento de transposição retire o alelo C_t e se desloque para uma nova localização, o alelo *C* pode ser funcional novamente: todas as células derivadas desta na qual ocorreu esse evento terão o genótipo *Cc* e serão púrpura. Essas células pigmentadas, circundadas por células incolores (C_tc), produzem uma mancha ou linha (chamada de setor) púrpura no grão amarelo (**Figura 18.31 D**). O tamanho do setor varia, dependendo de quando ocorreu a retirada do elemento de transposição do alelo C_t. Se for no início do desenvolvimento, então muitas células terão o alelo *C*

funcional e o setor pigmentado será grande; se for em uma etapa posterior do desenvolvimento, poucas células terão o alelo *C* funcional e o setor pigmentado será menor. Resolva o Problema 42

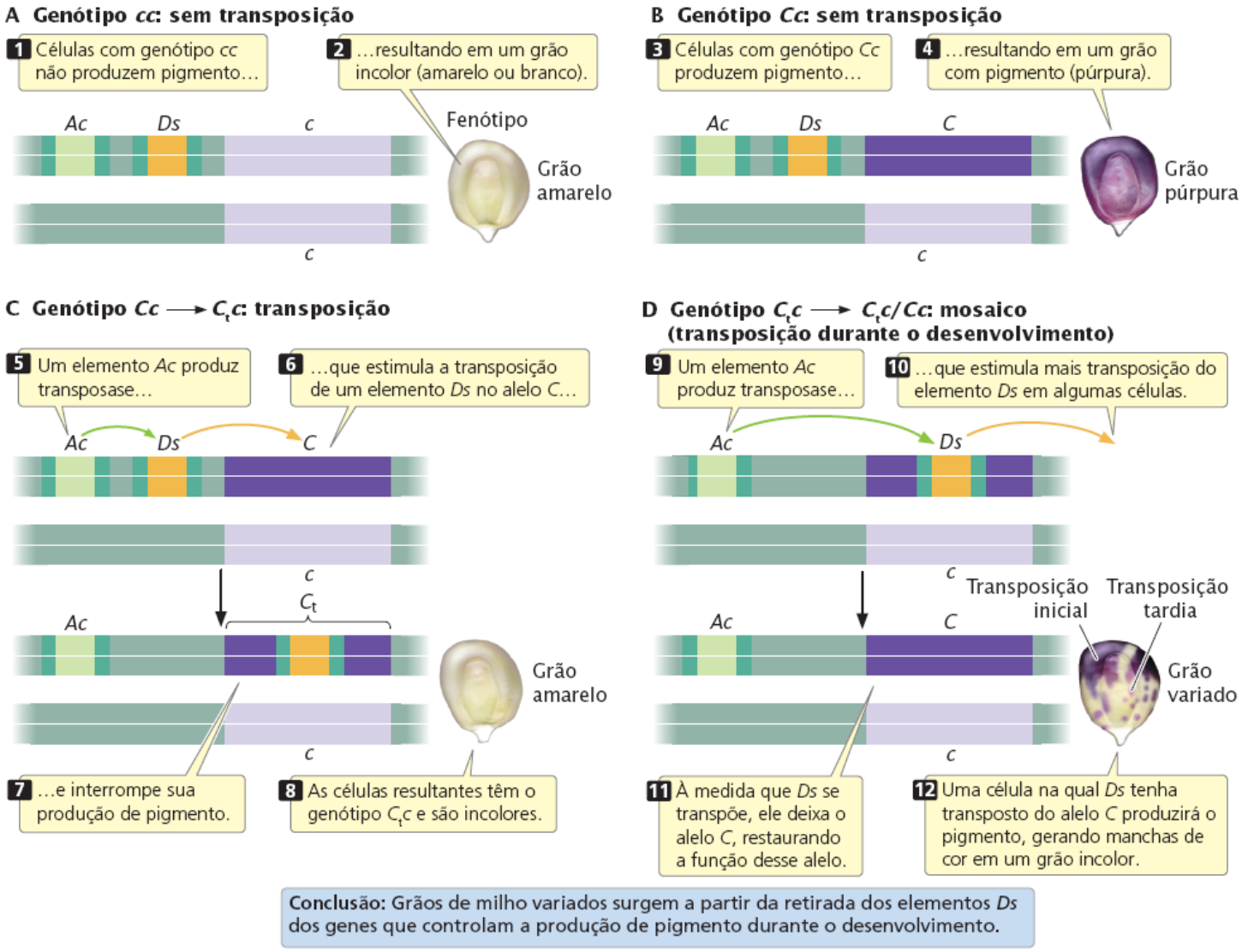


Figura 18.31 A transposição resulta em grãos de milho variados.

Elementos de transposição em Drosophila. Foram encontrados vários elementos de transposição diferentes em *Drosophila*. Uma família de elementos de transposição de *Drosophila* inclui os elementos *P*. A maioria dos elementos *P* funcionais tem 2.900 pb de comprimento, embora também existam elementos *P* menores com deleções. Cada elemento *P* tem repetições invertidas terminais e repetições diretas flanqueadoras no sítio da inserção. Como os elementos de transposição nas bactérias, os elementos *P* são transpósons do DNA. Cada elemento codifica uma transposase e um repressor da transposição.

O papel desse repressor ao controlar a transposição é demonstrado de forma dramática na **disgenesia híbrida**, que é o surgimento repentino de numerosas mutações, aberrações cromossômicas e a esterilidade nos descendentes de um cruzamento entre uma mosca-macho P^+ (*com* elementos P) e uma mosca-fêmea P^- (*sem* elementos). O cruzamento recíproco entre uma fêmea P^+ e um macho P^- produz descendentes normais.

A disgenesia híbrida surge a partir de uma explosão de transposição quando os elementos *P* são introduzidos em uma célula que não os tem. Em uma célula que tem os elementos *P*, um repressor no citoplasma inibe a transposição. Quando uma fêmea P^+ produz óvulos, a proteína repressora é incorporada ao citoplasma do óvulo, que evita a transposição no embrião e o surgimento das mutações. Os descendentes são férteis como os adultos (**Figura 18.32 A**). Entretanto, uma fêmea P^- não produz a proteína repressora; então, não há proteína armazenada no citoplasma de seus óvulos. O espermatozoide tem pouca ou nenhuma proteína no citoplasma, então um macho P^+ não contribui com a proteína repressora para seus descendentes. Quando os óvulos de uma fêmea P^- são fertilizados por espermatozoide de um macho P^+, a ausência de repressão torna possível que os elementos *P* fornecidos pelo espermatozoide sofram rápida transposição no embrião, causando a disgenesia híbrida (**Figura 18.32 B**).

A disgenesia híbrida e os elementos *P* têm atraído a atenção dos geneticistas porque esses elementos podem ter surgido em populações de *Drosophila melanogaster* nos últimos 50 anos e podem ter importância na evolução da espécie. Outras espécies

de *Drosophila* não apresentam os elementos *P*, que também estão ausentes das cepas laboratoriais de *D. melanogaster* originalmente coletada do tipo selvagem antes de 1960. Entretanto, atualmente a maioria das populações selvagens da espécie tem esses elementos. As cepas laboratoriais coletadas durante os anos 1970 são misturadas: algumas têm os elementos *P*, outras não. Isso sugere que os elementos *P* surgiram em algum momento no século passado e se disseminaram rapidamente por todas as populações selvagens de *D. melanogaster*. Como os cruzamentos entre os machos com elementos *P* e as fêmeas sem esse elemento provocam esterilidade, os elementos *P* e os elementos de transposição semelhantes têm o potencial para servir como mecanismos de isolamento reprodutivo entre as populações e podem participar da especiação (ver *Capítulo 26*). Essas observações sustentam a ideia de que os elementos de transposição são importantes na evolução. **Resolva o Problema 38**

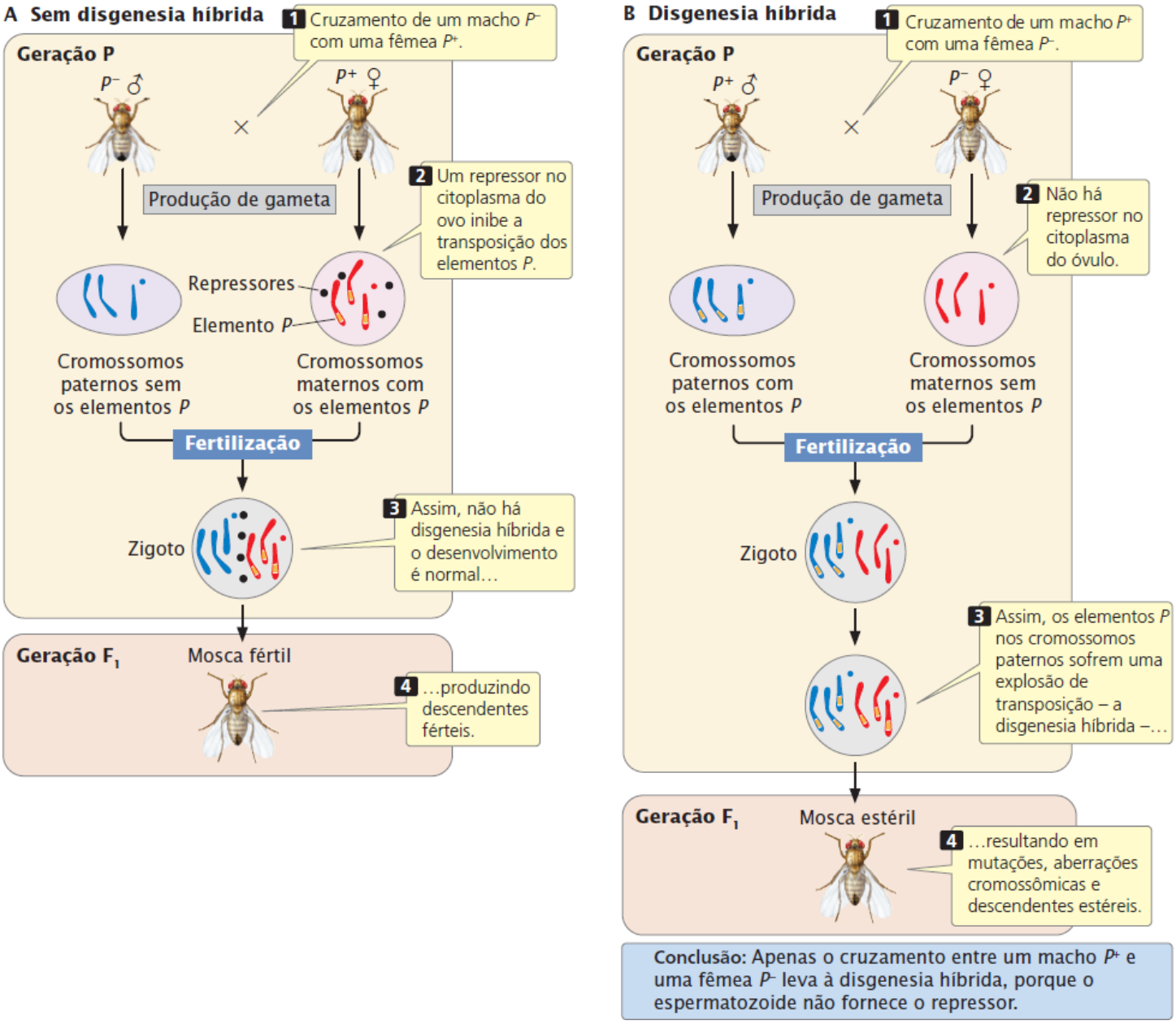


Figura 18.32 A disgenesia híbrida na *Drosophila* é causada pela transposição dos elementos *P*. (De W. Y. Chooi, *Genetics* 68:213-230, 1971.)

Elementos de transposição nos humanos. Um dos elementos de transposição mais comum no genoma humano é o *Alu*. Cada célula humana tem mais de 1 milhão de cópias semelhantes, mas não idênticas, de *Alu* em seus cromossomos. As sequências *Alu* são semelhantes ao gene que codifica a molécula 7S RNA, que transporta proteínas recém-sintetizadas pelo retículo endoplasmático. Essas sequências criam repetições diretas flanqueadoras quando se inserem no DNA e têm as características que sugerem que eles transpuseram por meio de um RNA intermediário.

O *Alu* pertence a uma classe de sequências repetitivas encontradas frequentemente nos mamíferos e em outros genomas. Essas sequências são coletivamente chamadas de elementos curtos dispersos (SINEs) e constituem cerca de 11% do genoma humano. A maioria dos SINEs é cópia de elementos de transposição que foram encurtados na extremidade 5′, provavelmente por causa do processo de transcrição reversa usado em sua transposição terminada antes que a sequência inteira fosse copiada.

Os SINEs foram identificados como a causa de mutações em mais de 20 casos de doença genética humana.

O genoma humano também tem muitos transpósons classificados como elementos longos dispersos (LINEs), que têm estrutura semelhante à dos retrovírus. Como os SINEs, a maioria dos LINEs no genoma humano foi encurtada na extremidade 5′. Os LINEs mais longos têm cerca de 6.000 pb, mas como a maioria das cópias é encurtada, a média dos LINEs é de apenas 900 pb. Existem aproximadamente 900.000 cópias de LINEs no genoma humano, constituindo coletivamente 21% do DNA humano total.

Conceitos

Existe uma grande variedade de elementos de transposição nos eucariotos. Alguns lembram elementos de transposição nos procariotos, com repetições invertidas terminais, e se transpõem como o DNA. Outros são retrotranspósons com repetições diretas longas nas suas extremidades e se transpõem por meio de um RNA intermediário.

Checagem dos conceitos 8

A disgenesia híbrida surge quando

a. uma mosca-macho com elementos *P* (P^+) cruza com uma mosca-fêmea sem os elementos *P* (P^-).
b. um macho P^- cruza com uma fêmea P^+.
c. um macho P^+ cruza com uma fêmea P^-.
d. um macho P^- cruza com uma fêmea P^-.

Conceitos conectantes

Tipos de elementos de transposição

Agora que examinamos alguns exemplos dos elementos de transposição, vamos revisar seus tipos principais (**Quadro 18.4**).

Os elementos de transposição podem ser divididos em duas classes principais com base na estrutura e no movimento. A Classe I inclui os retrotranspósons, que têm repetições diretas terminais e se transpõem usando RNA intermediários. Eles geram repetições diretas flanqueadoras nos seus pontos de inserção quando se transpõem no DNA. Os retrotranspósons não codificam a transposase, mas alguns tipos têm semelhanças na estrutura do retrovírus e carreiam sequências que produzem a transcriptase reversa. A transposição ocorre quando a transcrição produz um RNA intermediário, que é então transcrito em DNA pela transcriptase reversa e inserido no sítio-alvo. Os exemplos de retrotranspósons incluem elementos *Ty* na levedura e sequências *Alu* nos humanos. Os retrotranspósons não são encontrados nos procariotos.

A Classe II é composta por transpósons de DNA que têm repetições invertidas terminais e se transpõem como DNA. Como os transpósons da Classe I, eles todos geram repetições diretas flanqueadoras em seus pontos de inserção no DNA. Diferente dos transpósons da Classe I, todas as formas ativas dos elementos de transposição da Classe II codificam a transposase, que é necessária para seu movimento. Alguns também codificam resolvase, repressores e outras proteínas. Sua transposição pode ser replicativa ou não replicativa, mas eles nunca usam RNA intermediários. Os exemplos de elementos de transposição nessa classe incluem as sequências de inserção e todos os transpósons complexos nas bactérias, os elementos *Ac* e *Ds* no milho e os elementos *P* na *Drosophila*.

Elementos de transposição são importantes para a evolução humana

Os elementos de transposição exercem um papel importante na modelagem dos genomas de muitos organismos. Boa parte da grande variação no tamanho do genoma encontrado entre os organismos eucarióticos é devida a diferenças no número de elementos de transposição. Aproximadamente 45% do genoma humano é composto por remanescentes dos elementos de transposição e cerca de 50% de todas as mutações espontâneas em *Drosophila* são causadas pela transposição. A recombinação homóloga entre as cópias dos elementos de transposição é uma força importante na produção de duplicações de genes e outros rearranjos cromossômicos. Além disso, alguns elementos de transposição podem carrear DNA extra com eles quando se transpõem para um novo sítio, sendo o potencial para deslocar as sequências de DNA que regulam os genes para novos sítios, onde eles podem alterar a expressão dos genes.

Quadro 18.4 Características das duas principais classes de elementos genéticos de transposição.

	Estrutura	Genes codificados	Transposição	Exemplos
Classe I (retrotranspóson)	Repetições diretas longas e terminais; repetições diretas flanqueadoras curtas no sítio-alvo	Gene da transcriptase reversa (e às vezes outros)	Pelo RNA intermediário	*Ty* (levedura) *copia* (*Drosophila*) *Alu* (humano)
Classe II	Repetições invertidas terminais curtas; repetições diretas flanqueadoras curtas no sítio-alvo	Gene da transposase (e às vezes outros)	Através do DNA (replicativo e não replicativo)	*IS1* (*E. coli*) *Tn3* (*E. coli*) *Ac*, *Ds* (milho) Elementos *P* (*Drosophila*)

Elementos de transposição como parasitas genômicos. Como observamos, muitos elementos de transposição deixam uma cópia para trás quando se transpõem para uma nova localização (transposição copia e cola) e, portanto, aumentam em número em um genoma com o passar do tempo. Essa capacidade de se replicar e se disseminar significa que muitos elementos de transposição podem não ter finalidade para a célula; eles existem simplesmente porque são capazes de se replicar e se disseminar. A inserção dos elementos de transposição em um gene destruirá sua função, com consequências prejudiciais para a célula. Além disso, é provável que o tempo e a energia necessários para replicar grande número desses elementos seriam uma sobrecarga metabólica para a célula. Assim, esses elementos podem ser considerados parasitas genômicos que não fornecem benefício para a célula e podem até ser prejudiciais.

Domesticação dos elementos de transposição. Embora muitos desses elementos possam ser parasitas genômicos, alguns se desenvolveram para fins úteis para suas células hospedeiras. Esses transpósons são chamados de domesticados, o que indica que sua tendência de parasita foi substituída por propriedades úteis para a célula. Por exemplo, é provável que o mecanismo que gera a diversidade metabólica no sistema imunológico dos vertebrados (ver *Capítulo 22*) tenha se desenvolvido a partir de um elemento de transposição. As células imunológicas chamadas linfócitos têm a capacidade de unir vários segmentos de DNA que codificam as proteínas que reconhecem os antígenos. Esse mecanismo pode ter surgido a partir de um elemento de transposição que se inseriu na linhagem germinativa de um ancestral vertebrado há 450 milhões de anos.

Os elementos de transposição também participam da evolução do milho. O milho moderno foi domesticado a partir do teosinto na América Central há mais 8.000 anos. Uma das diferenças genéticas importantes entre o teosinto e o milho envolve um gene chamado *tb1*, que codifica um regulador transcricional responsável pela repressão do crescimento dos ramos laterais. No milho, a transcrição de *tb1* é elevada em comparação com a de teosinto, o que faz com que o milho moderno seja mais ereto e menos ramificado que o teosinto. Pesquisa recente demonstrou que a transcrição elevada de *tb1* no milho é o resultado de um elemento de transposição chamado *Hopscotch*, que se inseriu nas sequências reguladoras que controlam a transcrição de *tb1*. É provável que essa inserção tenha sido encontrada no teosinto como uma variante e foi selecionada pelos humanos durante a domesticação do milho porque ele produzia um formato de planta mais desejável. Os elementos de transposição também têm um papel na especiação, o processo pelo qual novas espécies surgem (ver a seção anterior sobre disgenesia híbrida).

Conceitos

Muitos elementos de transposição parecem ser parasitas genômicos, existindo em grande quantidade graças à capacidade de aumentar de forma eficiente suas cópias. O aumento no número de cópias dos elementos de transposição contribuiu para o grande tamanho de muitos genomas eucarióticos. Em muitos casos, os elementos de transposição foram adotados para funções celulares específicas.

18.5 Várias vias reparam as mudanças no DNA

A integridade do DNA está sob constante ataque de radiação, mutágenos químicos e mudanças que surgem espontaneamente. Apesar desses agentes prejudiciais, a taxa de mutação permanece incrivelmente baixa, graças à eficiência do sistema de reparo

do DNA.

Existem diversas vias complexas para reparar o DNA, mas podem ser feitas várias afirmações gerais sobre esse processo. Primeiro, a maioria dos mecanismos de reparo do DNA requer duas fitas de nucleotídios de DNA, porque a maioria substitui tais nucleotídios e é necessária uma fita molde para especificar a sequência de bases.

Uma segunda característica geral do reparo do DNA é a redundância, o que significa que muitos tipos de dano ao DNA podem ser corrigidos por mais de uma via de reparo. Essa redundância ilustra a extrema importância do reparo do DNA para a sobrevivência da célula: se um erro escapar do sistema de reparo, é provável que ele seja reparado por outro sistema, garantindo que quase todos os erros sejam corrigidos.

Vamos analisar os vários mecanismos gerais de reparo do DNA, reparo do pareamento errado, reparo direto, reparo por excisão de base, reparo por excisão de nucleotídio e reparo de quebras de fita dupla.

Reparo de pareamento errado

A replicação é um processo muito preciso: cada nova cópia de DNA tem menos de um erro por um bilhão de nucleotídios. Entretanto, no processo de replicação, as bases com pareamento errado são incorporadas no novo DNA com uma frequência de 10^{-4} a 10^{-5}; então, a maioria dos erros que surge inicialmente é corrigida e nunca se torna mutações permanentes. Algumas dessas correções são feitas na revisão (ver *Capítulo 12*) pelas DNA polimerases.

Muitos nucleotídios incorretamente inseridos que escapam da detecção da revisão são corrigidos pelo *reparo de pareamento errado*. As bases incorretamente pareadas são detectadas e corrigidas pelas enzimas de reparo de pareamento errado. Além disso, o sistema de reparo de pareamento errado corrige pequenas alças não pareadas no DNA, como as causadas pelo deslize de fita na replicação (ver Figura 18.11). Algumas repetições de nucleotídios podem formar estruturas secundárias na fita não pareada (ver Figura 18.4), tornando possível que elas escapem da detecção desse sistema.

Após a incorporação de erro ser identificada, as enzimas desse sistema de reparo retiram a seção da fita recém-sintetizada e preenchem o espaço com novos nucleotídios ao usar a fita de DNA original como molde. Para que essa estratégia funcione, o reparo de pareamento errado deve ter uma forma para diferenciar entre as fitas antiga e nova de DNA de modo que o erro de incorporação, mas não parte da fita original, seja removido.

As proteínas que fazem o reparo de pareamento errado na *E. coli* diferenciam entre as fitas antiga e nova por grupos metila em sequências especiais da fita antiga. Após a replicação, os nucleotídios adenina na sequência GATC são metilados. O processo de metilação é atrasado, e então, imediatamente após a replicação, a fita antiga está metilada e a nova, não (**Figura 18.33 A**). O complexo de reparo de pareamento errado produz uma sequência GATC não metilada próxima das bases malpareadas. Ele corta a fita não metilada no sítio GATC (**Figura 18.33 B**) e degrada a fita entre o corte e as bases malpareadas (**Figura 18.33 C**). A DNA polimerase e a DNA ligase preenchem o espaço da fita não metilada com nucleotídios corretamente pareados (**Figura 18.33 D**).

O reparo de pareamento errado nas células eucarióticas é semelhante àquele na *E. coli*, mas ainda não sabemos como as fitas antiga e nova são reconhecidas. Em alguns eucariotos, como as leveduras e as moscas-da-fruta, não há metilação detectável do DNA, e esse reparo é feito mesmo assim. Os humanos com mutações nos genes desse reparo exibem elevadas mutações somáticas e estão mais suscetíveis ao câncer de cólon.

Conceitos

As bases malpareadas e outras lesões de DNA são corrigidas pelo reparo de pareamento errado. As enzimas cortam uma seção da fita de DNA recém-sintetizada e a substituem por novos nucleotídios.

Checagem dos conceitos 9

O reparo de pareamento errado na *E. coli* diferencia as fitas antiga e nova do DNA com base:

a. nas diferenças na composição de base das duas fitas.
b. na modificação das proteínas histona.
c. nos análogos de base na fita nova.
d. nos grupos metila na fita antiga.

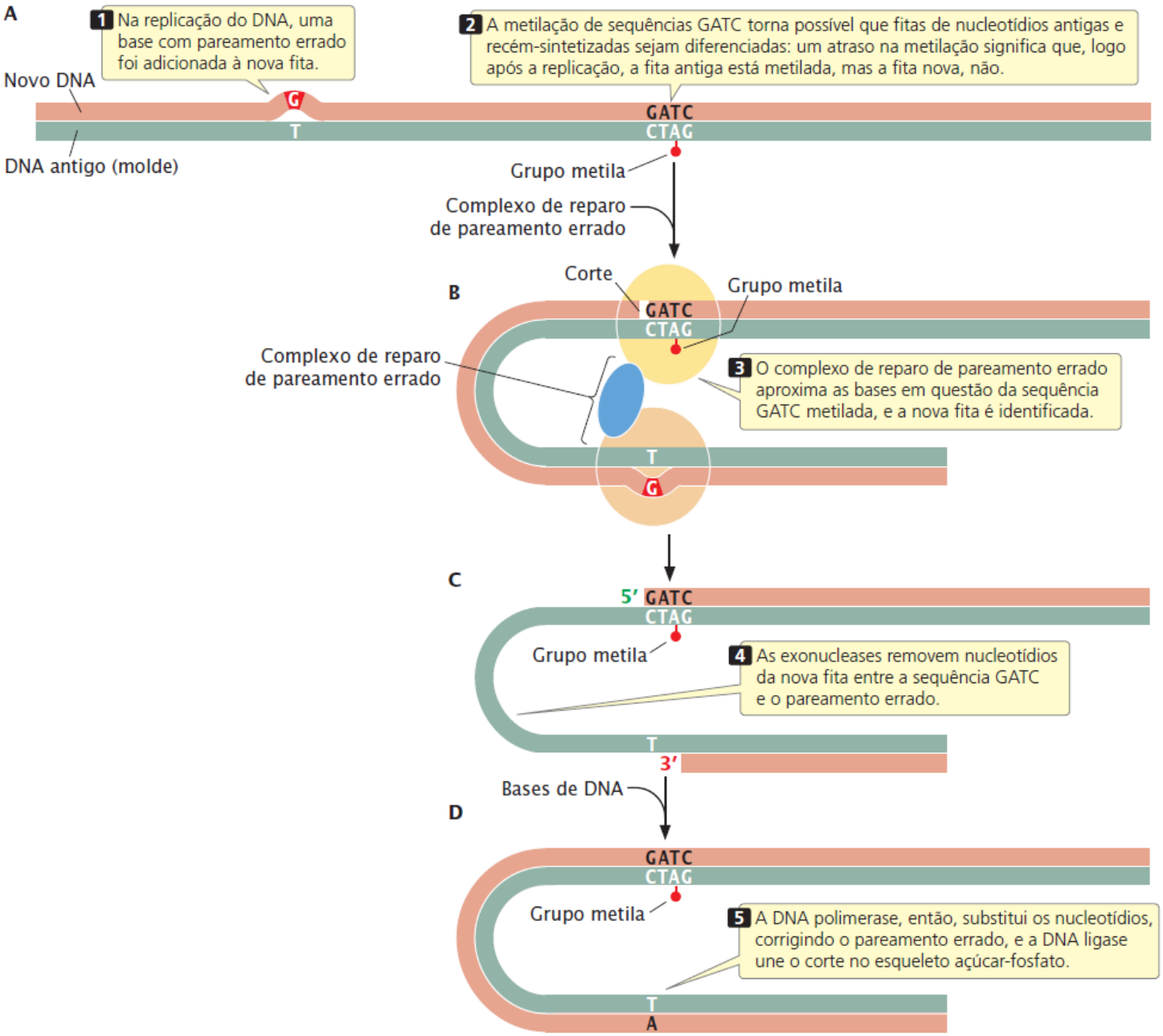


Figura 18.33 Muitos nucleotídios incorretamente inseridos que escapam da revisão são corrigidos pelo reparo de pareamento errado.

Reparo direto

O **reparo direto** não substitui os nucleotídios; ao contrário, ele os coloca em suas estruturas originais (corretas). Um dos mecanismos de reparo direto mais estudado é a fotorreativação dos dímeros de pirimidina induzidos pelo UV (ver Figura 18.20). A *E. coli* e algumas células eucarióticas têm uma enzima chamada fotoliase, que usa a energia capturada da luz para quebrar as ligações covalentes que unem as pirimidinas como um dímero.

O reparo direto também corrige a O^6-metilguanina, um produto da alquilação que pareia com adenina, produzindo as transversões G • C → T • A. Uma enzima chamada O^6-metilguanina-DNA metiltransferase remove o grupo metila da O^6-metilguanina, restaurando a base para guanina (**Figura 18.34**).

Reparo por excisão de bases

No **reparo por excisão de bases**, uma base modificada é primeiro retirada e então o nucleotídio inteiro é substituído. A excisão das bases modificadas é catalisada por um conjunto de enzimas chamado DNA glicosilases; cada uma reconhece e remove um tipo específico de base modificada ao romper a ligação que une aquela base ao átomo de carbono 1′ do açúcar desoxirribose (**Figura 18.35 A**). A uracila glicosilase, por exemplo, reconhece e remove a uracila produzida pela desaminação da citosina. Outras glicosilases identificam hipoxantina, 3-metiladenina, 7-metilguanina e outras bases modificadas.

Após a base ser removida, uma enzima chamada AP (apurínica ou apirimidínica) endonuclease corta a ligação fosfodiéster, e outras enzimas removem o açúcar desoxirribose (**Figura 18.35 B**). A DNA polimerase adiciona um ou mais novos nucleotídios

ao grupo 3′-OH exposto (**Figura 18.35 C**), substituindo a seção de nucleotídios da fita danificada. O corte no esqueleto fosfodiéster é selado pela DNA ligase (**Figura 18.35 D**), e a sequência original intacta é restaurada (**Figura 18.35 E**).

As bactérias usam DNA polimerase I para substituir os nucleotídios retirados, mas os eucariotos usam a DNA polimerase β, que não faz a revisão e tende a cometer erros. Em geral, a DNA polimerase β comete um erro a cada 4.000 nucleotídios inseridos. Cerca de 20.000 a 40.000 modificações de base por dia são reparadas por excisão da base, e então a DNA polimerase β pode introduzir cerca de 10 mutações por dia no genoma humano. Como esses erros são corrigidos? Pesquisa recente mostra que algumas AP endonucleases podem fazer revisão. Quando a DNA polimerase β insere um nucleotídio com a base errada no DNA, a DNA ligase não pode selar o corte no esqueleto açúcar-fosfato, pois os grupos 3′-OH e 5′-OH dos nucleotídios adjacentes não estão na orientação correta para a ligase conectá-los. Nesse caso, a AP endonuclease 1 detecta o pareamento errado e usa sua atividade de 3′ → 5′ exonuclease para retirar a base incorretamente pareada. A DNA polimerase β então usa sua atividade polimerase para preencher o nucleotídio ausente. Desse modo, a fidelidade do reparo por excisão de bases é mantida.

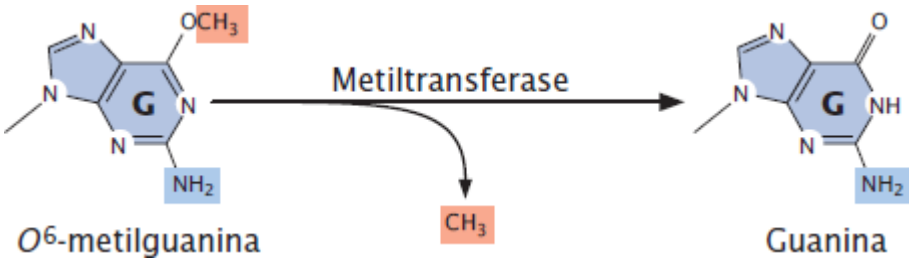
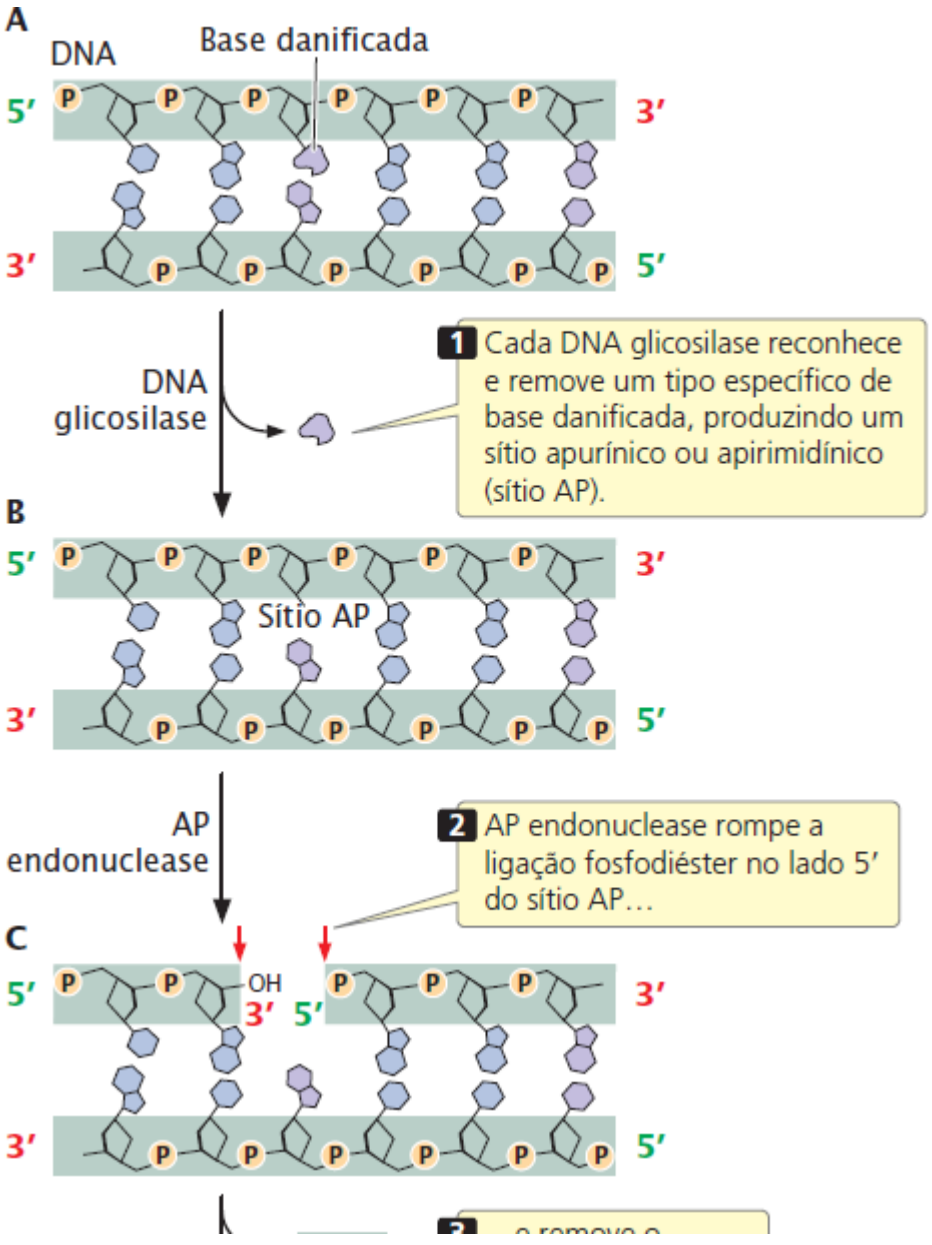


Figura 18.34 O reparo direto muda os nucleotídios de volta para suas estruturas originais.



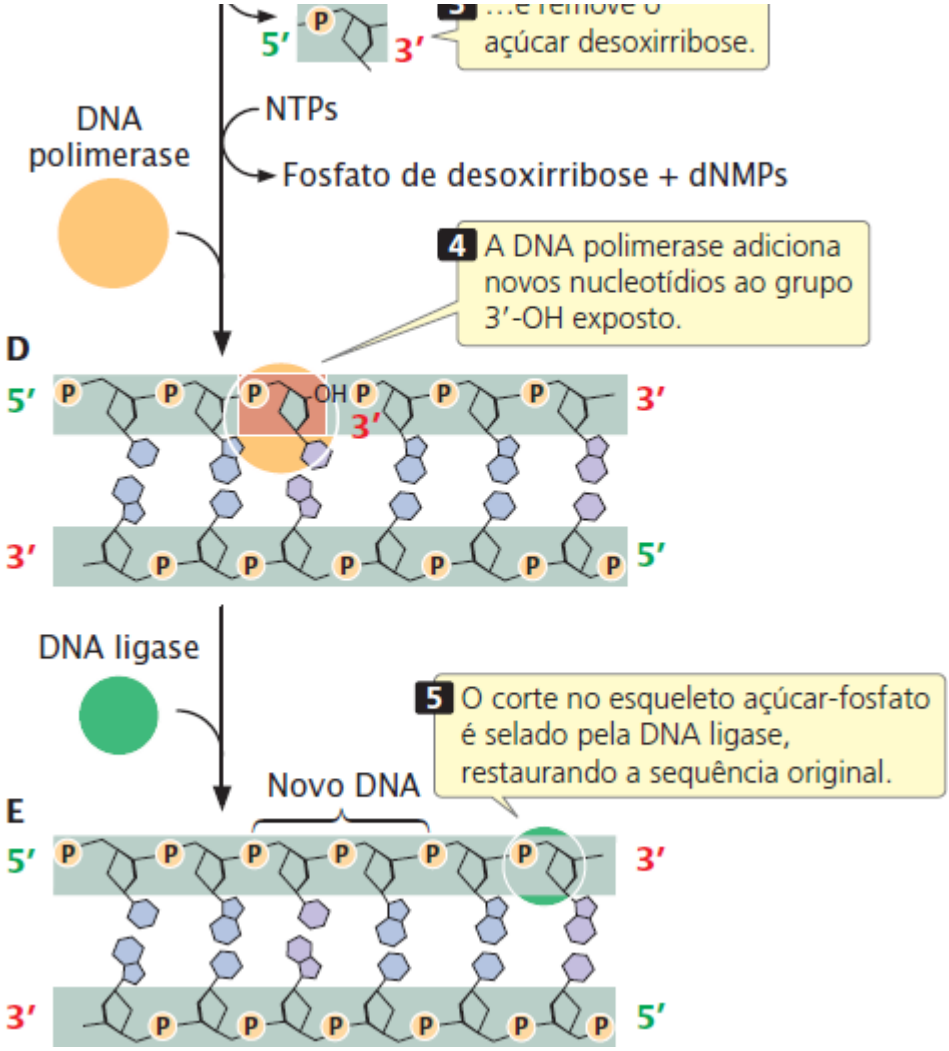


Figura 18.35 **O reparo por excisão de bases retira as bases modificadas e substitui um ou mais nucleotídios.**

Conceitos

Os mecanismos de reparo direto modificam nucleotídios alterados de volta para suas estruturas corretas. No reparo por excisão de bases, as enzimas glicosilase reconhecem e removem tipos específicos de bases modificadas. O nucleotídio inteiro é então removido, e uma seção da fita de polinucleotídios é substituída.

Checagem dos conceitos 10

Como os mecanismos de reparo direto diferem do reparo de pareamento errado e do reparo por excisão de bases?

Reparo por excisão de nucleotídio

Outra via de reparo é o **reparo por excisão de nucleotídio**, que remove lesões volumosas de DNA (como dímeros de pirimidina) que distorcem a dupla-hélice. O reparo por excisão de nucleotídio pode reparar muitos tipos diferentes de dano ao DNA e é encontrado nas células de todos os organismos, desde as bactérias até os humanos.

O processo de excisão de nucleotídio é complexo, vários genes participam dessa etapa nos humanos. Primeiro, um complexo de enzimas analisa o DNA, procurando distorções da sua configuração tridimensional (**Figura 18.36 A** e **B**). Quando uma distorção é detectada, enzimas adicionais separam as duas fitas de nucleotídios da região danificada, e as proteínas ligadoras de fita única estabilizam as fitas separadas (**Figura 18.36 C**). A seguir, um esqueleto de açúcar-fosfato da fita danificada é clivado em ambos os lados da lesão (**Figura 18.36 D**). Parte da fita danificada é removida pelas enzimas helicases (**Figura 18.36 E**), e a lacuna é preenchida pela DNA polimerase e "selada" pela DNA ligase (**Figura 18.36 F**).

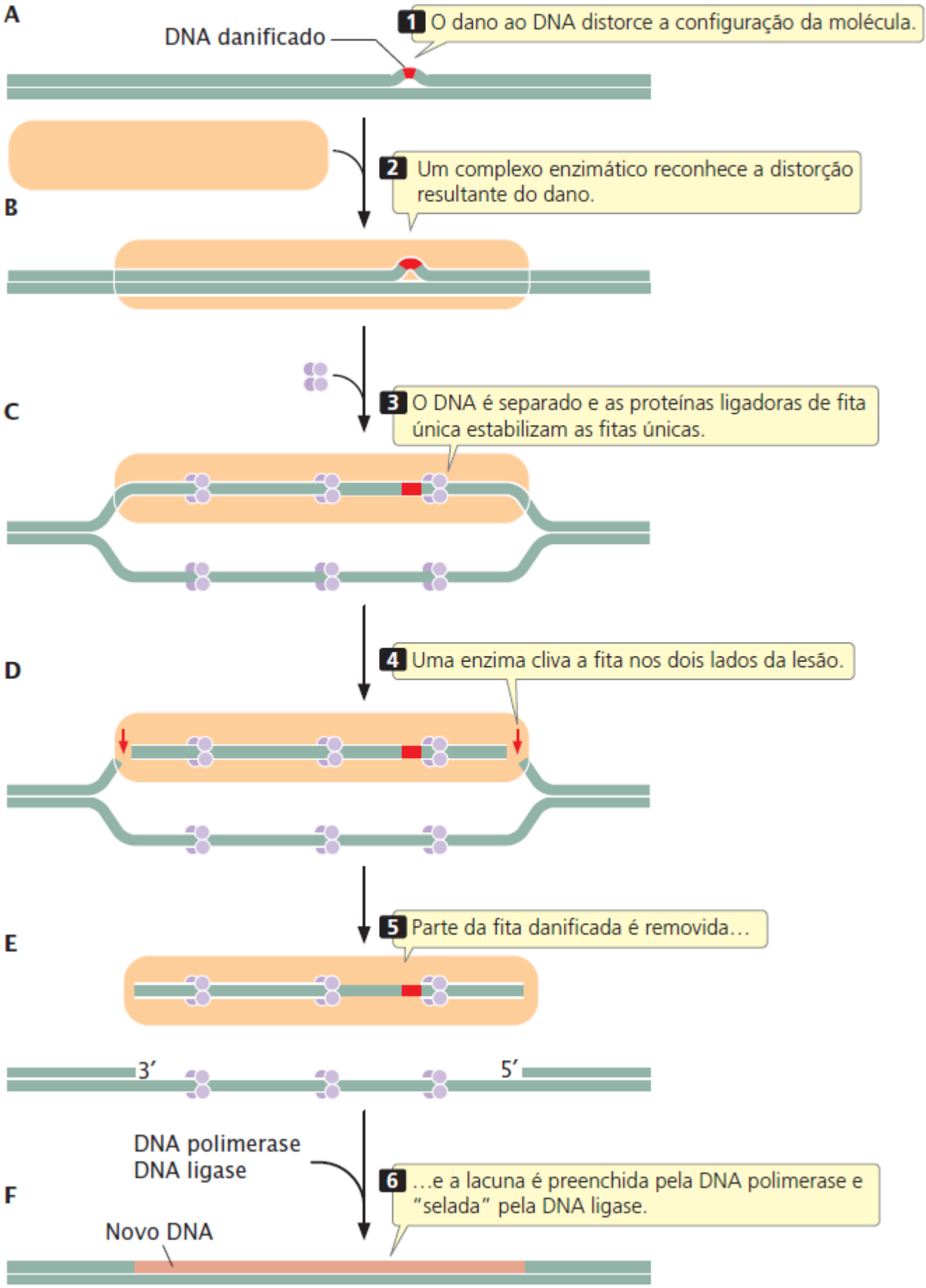


Figura 18.36 O reparo por excisão de nucleotídio remove lesões volumosas do DNA que distorcem a dupla-hélice.

Conceitos

O reparo por excisão de nucleotídio remove e substitui muitos tipos de DNA danificado que distorcem a estrutura do DNA. As duas fitas de DNA são separadas, uma seção do DNA contendo a distorção é

removida, a DNA polimerase preenche o espaço e a DNA ligase sela o espaço preenchido.

Conceitos conectantes

Via básica de reparo do DNA

Acabamos de examinar diferentes mecanismos de reparo do DNA. O que esses métodos têm em comum? E quais são suas diferenças? A maioria dos métodos de reparo do DNA depende de duas fitas, para que os nucleotídios na área danificada sejam removidos e substituídos. Os nucleotídios são substituídos no reparo de pareamento errado, reparo por excisão de bases e reparo por excisão de nucleotídio, mas não são substituídos por mecanismos de reparo direto.

Os mecanismos de reparo que incluem a remoção dos nucleotídios usam uma via comum de quatro etapas:

1. Detecção: a seção danificada do DNA é reconhecida.
2. Excisão: as endonucleases de reparo do DNA cortam o esqueleto fosfodiéster em um ou ambos os lados do DNA danificado e um ou mais nucleotídios são removidos.
3. Polimerização: a DNA polimerase adiciona nucleotídios ao grupo 3′-OH recém-exposto ao usar a outra fita como um molde e substituindo os nucleotídios danificados (e frequentemente alguns não danificados).
4. Ligação: a DNA ligase sela os cortes no esqueleto açúcar-fosfato.

As principais diferenças nos mecanismos de reparo de pareamento errado, excisão de bases e excisão de nucleotídios são os detalhes de detecção e excisão. No reparo por excisão de bases e pareamento errado, um único corte é feito no esqueleto açúcar-fosfato em um lado do dano; no reparo por excisão de nucleotídio, os cortes são feitos em ambos os lados da lesão do DNA. No reparo por excisão de bases, a DNA polimerase desloca os nucleotídios antigos à medida que adiciona novos nucleotídios à extremidade 3′ do corte; no reparo de pareamento errado, os nucleotídios antigos são degradados; e no reparo por excisão de nucleotídio, os nucleotídios são deslocados pelas enzimas helicases. Todos os três mecanismos usam DNA polimerase e ligase para preencher o espaço produzido pela excisão e remoção dos nucleotídios danificados.

Reparo de quebras de fita dupla

Um tipo comum de dano ao DNA é a quebra de fita dupla, na qual ambas as fitas da hélice do DNA são rompidas. As quebras de fita dupla são causadas por radiação ionizante, radicais livres oxidantes e outros agentes que danificam o DNA. Esses tipos de quebras são particularmente prejudiciais para a célula porque elas interrompem a replicação do DNA e podem levar a rearranjos cromossômicos, como deleções, duplicações, inversões e translocações. Existem duas vias principais para reparar as quebras de fita dupla: a recombinação homóloga e a junção de extremidades não homólogas.

Recombinação homóloga. A recombinação homóloga repara uma molécula de DNA rompida ao usar a informação genética idêntica ou quase idêntica contida em outra molécula de DNA, em geral uma cromátide-irmã – o mesmo mecanismo empregado no processo de recombinação homóloga responsável pelo *crossing over* (ver *Capítulo 12*). Essa recombinação inicia com a remoção de alguns nucleotídios nas extremidades rompidas, seguida por invasão, deslocamento e replicação da fita (ver Figura 12.20). Muitas das mesmas enzimas que fazem o *crossing over* são usadas no reparo de quebras de fita dupla pela recombinação homóloga: essas duas enzimas são BRCA1 e BRCA2. Os genes que codificam essas proteínas frequentemente sofrem mutação nas células do câncer de mama.

Junção de extremidades não homólogas. A junção de extremidades não homólogas repara as quebras de fita dupla usando um molde homólogo. Essa via é usada quando a célula está em G_1 e uma cromátide-irmã não está disponível para reparo pela recombinação homóloga. A junção de extremidades não homólogas usa proteínas que reconhecem as extremidades rompidas do DNA, liga as extremidades e então as une. Essa junção está mais suscetível a erros do que a recombinação homóloga e leva a deleções, inserções e translocações. O **Quadro 18.5** resume os diferentes tipos de reparo do DNA.

DNA polimerases translesão

Como foi discutido no Capítulo 12, as DNA polimerases de alta fidelidade que normalmente fazem a replicação operam em alta velocidade, como um trem de alta velocidade, exigindo um trilho liso – um molde não distorcido. Algumas mutações, como os dímeros de pirimidina, produzem distorções na estrutura tridimensional da hélice do DNA, bloqueando a replicação das

polimerases de alta velocidade. Quando são encontradas as distorções no molde, DNA polimerases translesão especializadas assumem a replicação e contornam as lesões.

As polimerases translesão contornam as lesões volumosas, mas produzem erros no processo. Assim, as polimerases translesão tornam possível que a replicação continue ao custo de introduzir mutações na sequência. Algumas dessas mutações são corrigidas por sistemas de reparo do DNA, mas outras escapam da detecção.

Um exemplo de DNA polimerase translesão é a polimerase η (eta), que contorna os dímeros de pirimidina nos eucariotos. A polimerase η insere AA oposto a um dímero de piridimina. Essa estratégia parece ser viável porque cerca de dois terços dos dímeros de pirimidina são dímeros de timina. Entretanto, a inserção de AA oposta ao dímero CT resulta em uma transversão C • G → T • A. Dessa maneira, a polimerase η tende a introduzir mutações na sequência de DNA.

Quadro 18.5 Resumo dos mecanismos comuns de reparo do DNA.

Sistema de reparo	Tipo de dano reparado
Pareamento errado	Erros de replicação, incluindo bases malpareadas e deslizamento da fita
Direto	Dímeros de pirimidina; outros tipos específicos de alterações
Excisão de base	Bases anormais, bases modificadas e dímeros de pirimidina
Excisão de nucleotídio	Dano ao DNA que distorce a dupla-hélice, incluindo bases anormais, bases modificadas e dímeros de pirimidina
Recombinação homóloga	Quebras de fita dupla
Junção de extremidades não homólogas	Quebras de fita dupla

Conceitos

Existem duas vias principais para o reparo de quebras de fita dupla no DNA: a recombinação homóloga e a junção de extremidades não homólogas. As DNA polimerases translesão especiais tornam possível que a replicação continue ao contornar distorções volumosas no DNA, mas introduzem erros à medida que contornam a região distorcida.

Doenças genéticas e reparo do DNA defeituoso

Várias doenças humanas estão conectadas a defeitos no reparo do DNA. Essas doenças estão associadas a altas incidências de cânceres específicos, porque os defeitos no reparo do DNA levam a maiores taxas de mutação. Esse conceito é discutido com mais detalhes no Capítulo 23.

Entre as doenças humanas de reparo do DNA mais estudadas está o xeroderma pigmentoso, uma rara condição autossômica recessiva que inclui pigmentação anormal da pele e sensibilidade aguda à luz solar. As pessoas com essa doença também apresentam forte predisposição ao câncer de pele, com uma incidência variando de 1.000 a 2.000 vezes que a encontrada nas pessoas não afetadas.

A luz solar inclui um forte componente da UV, então a exposição à luz solar produz dímeros de pirimidina no DNA das células da pele. Embora as células humanas não tenham fotoliase (a enzima que repara os dímeros de pirimidina nas bactérias), a maioria dos dímeros de pirimidina nos humanos pode ser corrigida pelo reparo por excisão de nucleotídio (ver Figura 18.36). Entretanto, as células da maioria das pessoas com xeroderma pigmentoso apresentam defeito no reparo por excisão de nucleotídio, e muitos dos seus dímeros de pirimidina ficam sem correção e podem levar ao câncer.

O xeroderma pigmentoso pode surgir a partir de defeitos em vários genes diferentes. Algumas pessoas com xeroderma pigmentoso têm mutações em um gene que codifica a proteína que reconhece e se liga ao DNA danificado; outras têm mutações em um gene que codifica a helicase. Outras, ainda, apresentam defeitos nos genes que participam do corte da fita danificada nos lados 5′ e 3′ do dímero de pirimidina. Algumas pessoas têm uma forma ligeiramente diferente da doença (xeroderma

pigmentoso variante) devido a mutações no gene que codifica a polimerase η, a DNA polimerase translesão que contorna os dímeros de pirimidina.

Outra doença genética causada por um reparo defeituoso do DNA é uma forma herdada de câncer de cólon hereditária, o câncer de cólon não poliposo (HNPCC). Esse é um dos cânceres hereditários mais comuns, responsável por cerca de 15% dos cânceres de cólon. Os achados das pesquisas indicam que o HNPCC surge a partir de mutações nas proteínas responsáveis pelo reparo de pareamento errado (ver Figura 18.33). O **Quadro 18.6** resume algumas doenças genéticas associadas ao reparo defeituoso do DNA. Resolva o Problema 44

Conceitos

Os defeitos no reparo do DNA são a causa subjacente de várias doenças genéticas. Muitas dessas doenças são caracterizadas por predisposição ao câncer.

Checagem dos conceitos 11

Por que os defeitos no reparo do DNA estão frequentemente associados ao aumento no câncer?

Quadro 18.6 Doenças genéticas associadas aos defeitos nos sistemas de reparo do DNA.

Doença	Sintomas	Defeito genético
Xeroderma pigmentoso	Manchas sardentas na pele, sensibilidade à luz solar, predisposição ao câncer de pele	Defeitos no reparo por excisão de nucleotídio
Síndrome de Cockayne	Nanismo, sensibilidade à luz solar, envelhecimento prematuro, surdez, déficit intelectual	Defeitos no reparo por excisão de nucleotídio
Tricotiodistrofia	Fios de cabelo quebradiços, anomalias na pele, baixa estatura, desenvolvimento sexual imaturo, traços faciais característicos	Defeitos no reparo por excisão de nucleotídio
Câncer de cólon não poliposo hereditário	Predisposição ao câncer de cólon	Defeitos no reparo de pareamento errado
Anemia de Fanconi	Maior pigmentação da pele, anomalias do esqueleto, coração e rins, predisposição à leucemia	Possíveis defeitos no reparo de ligações cruzadas entre as fitas
Síndrome de Li-Fraumeni	Predisposição ao câncer em muitos tecidos diferentes	Defeitos na resposta ao dano ao DNA
Síndrome de Werner	Envelhecimento prematuro, predisposição ao câncer	Defeitos na recombinação homóloga

Resumo dos conceitos

- Mutações são mudanças herdadas na informação genética. As mutações somáticas ocorrem nas células somáticas; as mutações germinativas ocorrem em células que dão origem aos gametas
- O tipo mais simples de mutação é uma substituição de base, uma mudança em um único par de bases do DNA. As transições são substituições de base nas quais as purinas são substituídas por purinas ou pirimidinas são substituídas por pirimidinas. As transversões são substituições de base nas quais uma purina substitui uma pirimidina ou uma pirimidina substitui uma purina
- Inserções são a adição de nucleotídios e deleções são a remoção dos nucleotídios; essas mutações mudam o quadro de leitura

do gene

- As repetições expandidas de nucleotídios são mutações nas quais o número de cópias de um conjunto de nucleotídios aumenta com o tempo; elas são responsáveis por várias doenças genéticas humanas
- Uma mutação *missense* altera a sequência codificadora de modo que um aminoácido substitui outro. Uma mutação *nonsense* muda um códon que especifica um aminoácido em um códon de terminação. Uma mutação silenciosa produz um códon sinônimo que especifica o mesmo aminoácido que a sequência original, enquanto uma mutação neutra altera a sequência de aminoácidos, mas não muda o funcionamento da proteína. Uma mutação supressora reverte o efeito de uma mutação anterior em um sítio diferente e pode ser intragênica (dentro do mesmo gene que a mutação original) ou intergênica (dentro de um gene diferente)
- A taxa de mutação é a frequência na qual uma mutação específica surge em uma população. As taxas de mutação são influenciadas por fatores genéticos e ambientais
- Algumas mutações ocorrem espontaneamente. Essas mutações incluem o pareamento errado de bases na replicação e a depurinação e desaminação espontâneas
- Podem surgir inserções e deleções por causa do deslizamento da fita na replicação e do *crossing over* desigual
- Os análogos de base podem ser incorporados no DNA durante a replicação e pareiam com a base errada em subsequentes eventos de replicação. Os agentes alquilantes e a hidroxilamina modificam a estrutura química das bases e levam a mutações. Os agentes intercalantes se inserem na molécula de DNA e provocam adições e deleções de um nucleotídio. As reações oxidativas alteram as estruturas químicas das bases
- A radiação ionizante é mutagênica, alterando as estruturas das bases e rompendo as ligações fosfodiéster. A luz ultravioleta produz dímeros de pirimidina, que bloqueiam a replicação. As bactérias usam a resposta SOS para superar os bloqueios de replicação produzidos pelos dímeros de pirimidina e outras lesões no DNA
- O teste de Ames usa bactérias para avaliar o potencial mutagênico de substâncias químicas
- Os elementos de transposição são sequências de DNA móveis que se inserem em vários locais dentro de um genoma e provocam mutações e rearranjos de DNA
- A maioria dos elementos de transposição tem duas características em comum: repetições invertidas terminais e geração de repetições diretas curtas no DNA no ponto de inserção
- A transposição pode ocorrer por meio de uma molécula de DNA ou pela produção de uma molécula de RNA que é então transcrita de forma reversa em DNA. A transposição pode ser replicativa, na qual o elemento de transposição é copiado e a cópia se desloca para um novo sítio, ou não replicativa, na qual o elemento de transposição é cortado do sítio antigo e se desloca para um novo
- Os retrotranspósons se transpõem por meio de moléculas de RNA que sofrem transcrição reversa para produzir DNA
- As sequências de inserção são pequenos elementos de transposição bacterianos que carreiam apenas a informação necessária para seu próprio movimento. Os transpósons compostos nas bactérias são elementos compostos de DNA mais complexos entre duas sequências de inserção. Alguns elementos de transposição complexos nas bactérias não têm sequências de inserção
- Os transpósons de DNA nas células eucarióticas são semelhantes aos encontrados nas bactérias, terminando em repetições invertidas curtas e produzindo repetições diretas flanqueadoras no ponto de inserção. Outros são retrotranspósons, com estrutura semelhante aos retrovírus e se transpondo através de RNA intermediários
- Os transpósons tiveram um papel importante na evolução do genoma
- A maioria dos danos ao DNA é corrigida por mecanismos de reparo do DNA. Esses mecanismos incluem reparo de pareamento errado, reparo direto, reparo por excisão de bases, reparo por excisão de nucleotídio e outras vias de reparo. A maioria requer duas fitas de DNA e exibe alguma sobreposição nos tipos de dano que reparam
- As quebras de fita dupla são reparadas pela recombinação homóloga e junção de extremidades não homólogas. As DNA polimerases translesão especiais tornam possível que a replicação supere as distorções volumosas no DNA
- Os defeitos no reparo do DNA são a causa principal de várias doenças genéticas.

Termos importantes

Agente intercalante
Análogo da base
Crossing over desigual
Deleção
Deleção no quadro de leitura
Depurinação
Desaminação

Deslizamento da fita
Dímero de pirimidina
Disgenesia híbrida
Elemento de transposição
Erro de incorporação
Erro de replicação
Inserção
Inserção no quadro de leitura
Mutação
Mutação adaptativa
Mutação condicional
Mutação de ganho de função
Mutação de perda de função
Mutação direta
Mutação espontânea
Mutação gênica
Mutação germinativa
Mutação induzida
Mutação letal
Mutação *missense*
Mutação neutra
Mutação *nonsense*
Mutação no quadro de leitura (*frameshift*)
Mutação reversa (reversão)
Mutação silenciosa
Mutação somática
Mutação supressora
Mutação supressora intergênica
Mutação supressora intragênica
Mutágeno
Reparo direto
Reparo por excisão de bases
Reparo por excisão de nucleotídio
Repetição direta flanqueadora
Repetição expandida de nucleotídios
Repetição invertida terminal
Retrotranspóson
Sequência de inserção (IS)
Sistema SOS
Substituição de base
Taxa de mutação
Teste de Ames
Transição
Transposase
Transposição
Transposição não replicativa
Transposição replicativa
Transpóson composto
Transpóson de DNA
Transversão

Respostas da Checagem dos conceitos

1. c.
2. Uma mutação reversa restaura o fenótipo original ao mudar a sequência de DNA de volta para a sequência do tipo

selvagem. Uma mutação repressora restaura o fenótipo ao provocar uma mudança adicional ao DNA em um sítio que é diferente do sítio da mutação original.

3. A frequência com que surgem mudanças no DNA, como essas mudanças são reparadas pelos mecanismos de reparo do DNA e nossa capacidade para detectar a mutação.
4. c.
5. Na transposição, são feitos os cortes escalonados no DNA e o elemento de transposição se insere no corte. Posteriormente, a replicação de segmentos de fita única de DNA cria repetições diretas flanqueadoras curtas em cada lado do elemento de transposição inserido.
6. A transposição resulta em mutações, porque o elemento de transposição se insere em um gene, destruindo sua função. Os rearranjos cromossômicos surgem porque a transposição inclui a quebra e a troca de sequências de DNA. Além disso, múltiplas cópias de um elemento de transposição podem sofrer recombinação homóloga, produzindo rearranjos cromossômicos.
7. d.
8. a.
9. d.
10. Os mecanismos de reparo direto retornam uma base alterada para sua estrutura correta sem remover nem substituir os nucleotídios. O reparo de pareamento errado e o reparo por excisão de bases removem e substituem os nucleotídios.
11. Mudanças na estrutura do DNA podem não sofrer reparo em pessoas com defeitos nos mecanismos de reparo do DNA. Consequentemente, ocorre um número maior de mutações em todos os genes, incluindo os que predispõem para o câncer. Essa observação indica que o câncer surge a partir de mutações no DNA.

Problemas desenvolvidos

Problema 1

As mutações produzidas pelos seguintes compostos são revertidas pelas substâncias apresentadas. Quais conclusões você pode tirar sobre a natureza das mutações originalmente produzidas por esses compostos?

	Mutações produzidas pelo composto	Revertidas por			
		5-bromouracila	EMS	Hidroxilamina	Acridina laranja
a.	1	Sim	Sim	Não	Não
b.	2	Sim	Sim	Algumas	Não
c.	3	Não	Não	Não	Sim
d.	4	Sim	Sim	Sim	Sim

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
As conclusões que você pode tirar sobre os tipos de mutações produzidos por cada composto com base nas substâncias que revertem as mutações.

Quais informações são fornecidas para solucionar o problema?
Quais substâncias revertem as mutações produzidas por cada composto.

Para a solução deste problema, revise:
Mutações quimicamente induzidas, na Seção 18.2.

Etapas para a solução

a. Mutações produzidas pelo composto 1 são revertidas pela 5-bromouracila, que produz as transições A • T → G • C e G • C → A • T, o que nos diz que o composto 1 produz substituições de base única que podem incluir a geração dos pares A • T ou G • C. As mutações produzidas pelo composto 1 são revertidas pelo EMS, que, como a 5-bromouracila, produz

tanto a transição A • T → G • C quanto a transição G • C → A • T; então, nenhuma informação adicional é fornecida aqui. A hidroxilamina não reverte as mutações produzidas pelo composto 1. Como a hidroxilamina produz apenas as transições C • G → T • A, agora sabemos que o composto 1 não gera os pares de base C • G. Acridina laranja, um agente intercalante que produz mutações no quadro de leitura, também não reverte as mutações, revelando que o composto 1 produz apenas substituições de par de base único, não inserções nem deleções. Em resumo, o composto causa substituições de base única que geram os pares de base T • A, mas não G • C.

b. O composto 2 gera mutações que são revertidas por 5-bromouracila e EMS, indicando que ele pode produzir os pares de base G • C ou A • T. Algumas dessas mutações são revertidas por hidroxilamina, que produz apenas as transições C • G → T • A, indicando que algumas das mutações produzidas pelo composto 2 são pares de base C • G. Nenhuma das mutações é revertida pela acridina laranja; então, o composto 2 não induz inserções ou deleções. Em resumo, o composto 2 produz substituições de base única que geram os pares de base G • C e A • T.

Dica: A capacidade de vários compostos de produzir mutações reversas revela informações importantes sobre a natureza da mutação original.

c. O composto 3 produz mutações que são revertidas apenas pela acridina laranja; então, parece que o composto 3 produz apenas inserções e deleções.

d. A mutação causada pelo composto 4 é revertida pela 5-bromouracila, EMS, hidroxilamina e acridina laranja, indicando que esse composto produz substituições de base única, que incluem os pares de base G • C e A • T, inserções e deleções.

Problema 2

Algumas sequências repetidas nos eucariotos são flanqueadas por repetições diretas curtas, sugerindo que elas se originaram como elementos de transposição. Essas mesmas sequências não têm íntrons e apresentam uma tira de nucleotídios timina em uma das extremidades. Esses elementos se transpuseram por meio de sequências de DNA ou RNA? Explique sua justificativa.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

Se os elementos de transposição agem por meio de DNA ou RNA intermediário e por que você chegou a essa conclusão.

Quais informações são fornecidas para solucionar o problema?

- O elemento é flanqueado por repetições diretas curtas
- O elemento não tem íntrons, mas tem uma tira de Ts em uma extremidade.

Para a solução deste problema, revise:

Transposição, na Seção 18.4.

Etapas para a solução

A ausência de *íntrons* e a tira de nucleotídios timina (que seriam complementares aos nucleotídios adenina no RNA) em uma extremidade são características de RNA processado. Essas semelhanças ao RNA sugerem que o elemento foi originalmente transcrito em mRNA, processado para remover os íntrons e adicionar uma cauda poli(A), e então transcrito de forma reversa em um DNA complementar que foi inserido no cromossomo.

Lembrete: Pré-mRNAs nos eucariotos são processados: um cap 5′ é adicionados, íntrons são removidos e uma cauda (poli) 3′ é adicionada.

Questões de compreensão

Seção 18.1

1. Qual é a diferença entre uma transição e uma transversão? Que tipo de substituição de base é a mais comum?
2. Descreva brevemente o que são as repetições expandidas de nucleotídios. Como elas contribuem para o fenômeno da

antecipação?

3. Qual é a diferença entre uma mutação *missense* e uma mutação *nonsense*? Entre uma mutação silenciosa e uma mutação neutra?
4. Descreva brevemente duas formas diferentes nas quais os supressores intragênicos podem reverter os efeitos das mutações.

Seção 18.2

5. Como surgem as inserções e deleções?
6. Como os análogos de base levam a mutações?
7. Como os agentes alquilantes, ácido nitroso e hidroxilamina produzem mutações?

Seção 18.3

8. Qual é o objetivo do teste de Ames? Como as bactérias *his*$^-$ são usadas nesse teste?

Seção 18.4

9. Quais são as características gerais encontradas em muitos elementos de transposição?
10. Como um retrotranspóson se desloca?
11. Desenhe a estrutura de uma típica sequência de inserção e identifique suas partes.
12. Explique como os elementos *Ac* e *Ds* produzem grãos de milho variados.
13. Quais são as diferenças entre os elementos de transposição das classes I e II?
14. Por que os elementos de transposição são chamados de parasitas genômicos?

Seção 18.5

15. Liste pelo menos três diferentes tipos de reparo do DNA e explique resumidamente como cada um funciona.
16. Quais são os dois principais mecanismos para o reparo de quebras de fita dupla? Quais são suas diferenças?

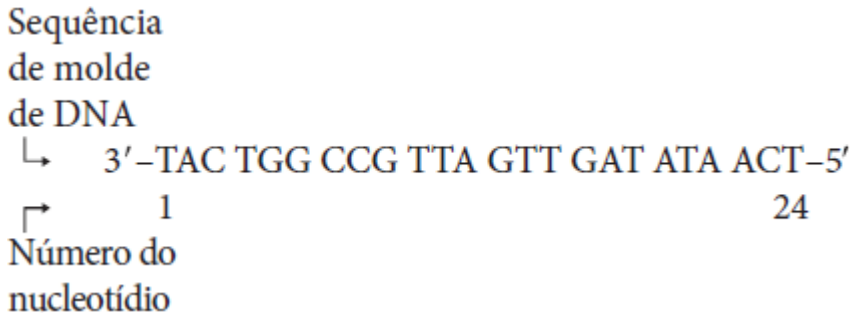
Questões e problemas aplicados

Seção 18.1

17. Um códon que especifica o aminoácido Gly sofre substituição de base única para se tornar uma mutação *nonsense*. De acordo com o código genético fornecido na Figura 15.8, essa mutação é uma transição ou uma transversão? Em qual posição do códon ocorre a mutação?

*18. a. Se ocorrer uma única transição em um códon que especifica Phe, quais aminoácidos poderiam ser especificados pela sequência da mutação?
    b. Se ocorrer uma transversão simples em um códon que especifica Phe, quais aminoácidos poderiam ser especificados pela sequência da mutação?
    c. Se ocorrer uma única transição em um códon que especifica Leu, quais aminoácidos podem ser especificados pela sequência da mutação?
    d. Se ocorrer uma única transversão em um códon que especifica Leu, quais aminoácidos podem ser especificados pela sequência da mutação?

19. A hemoglobina é uma proteína complexa com quatro cadeias de polipeptídios. A hemoglobina normal encontrada nos adultos – chamada de hemoglobina adulta – tem duas cadeias de polipeptídios alfa e duas cadeias beta, que são codificadas por diferentes *loci*. A hemoglobina falciforme, que causa a anemia falciforme, surge a partir de uma mutação na cadeia beta da hemoglobina adulta. A hemoglobina adulta e a hemoglobina falciforme têm um único aminoácido diferente: o sexto aminoácido de uma extremidade da hemoglobina adulta é o ácido glutâmico, enquanto a hemoglobina falciforme tem valina nessa mesma posição. Após consultar o código genético fornecido na Figura 15.8, indique o tipo e a localização da mutação que deu origem à anemia falciforme.
20. A sequência de nucleotídios a seguir é encontrada na fita molde de DNA. Primeiro, determine os aminoácidos da proteína codificada por essa sequência ao usar o código genético fornecido na Figura 15.8. Então, indique a sequência de aminoácidos alterada da proteína que será encontrada em cada uma das mutações a seguir:



a. Mutante 1: Uma transição no nucleotídio 11
b. Mutante 2: Uma transição no nucleotídio 13
c. Mutante 3: Uma deleção de um nucleotídio no nucleotídio 7
d. Mutante 4: Uma transversão A → T no nucleotídio 15
e. Mutante 5: Uma adição de TGG após o nucleotídio 6
f. Mutante 6: Uma transição no nucleotídio 9

21. Desenhe um grampo (*hairpin*) semelhante ao apresentado na Figura 18.4 para a sequência repetida encontrada na síndrome do X frágil (ver Quadro 18.1).

*22. Um polipeptídio tem a seguinte sequência de aminoácidos:

Met-Ser-Pro-Arg-Leu-Glu-Gly

A sequência de aminoácidos desse polipeptídio foi determinada em vários mutantes listados em partes *a* a *e*. Para cada mutante, indique o tipo de mutação que ocorreu no DNA (substituição de base única, inserção, deleção) e o efeito fenotípico da mutação (mutação *nonsense*, mutação *missense*, mutação no quadro de leitura etc.)

a. Mutante 1: Met-Ser-Ser-Arg-Leu-Glu-Gly
b. Mutante 2: Met-Ser-Pro
c. Mutante 3: Met-Ser-Pro-Asp-Trp-Arg-Asp-Lys
d. Mutante 4: Met-Ser-Pro-Glu-Gly
e. Mutante 5: Met-Ser-Pro-Arg-Leu-Leu-Glu-Gly

*23. Um gene codifica a proteína com a seguinte sequência de aminoácidos:

Met-Trp-His-Arg-Ala-Ser-Phe

Ocorre uma mutação no gene. Uma proteína mutante tem a seguinte sequência de aminoácidos:

Met-Trp-His-Ser-Ala-Ser-Phe

Um supressor intragênico restaura a sequência de aminoácidos da proteína original:

Met-Trp-His-Arg-Ala-Ser-Phe

Indique pelo menos um exemplo de mudança de base que poderia provocar a mutação original e o supressor intragênico. (Consulte o código genético na Figura 15.8.)

24. Um gene codifica a proteína com a seguinte sequência de aminoácidos:

Met-Lys-Ser-Pro-Ala-Thr-Pro

Ocorre uma mutação *nonsense* a partir da substituição de um par de base nesse gene, resultando em uma proteína com a sequência de aminoácidos Met-Lys. Uma mutação supressora intergênica possibilita que o gene produza a proteína de comprimento integral. Com a mutação original e o supressor intergênico, o gene agora produz uma proteína com a seguinte sequência de aminoácidos:

Met-Lys-Cys-Pro-Ala-Thr-Pro

Indique a localização e a natureza da mutação original e da mutação supressora intergênica.

Seção 18.2

*25. As mutações *nonsense* podem ser revertidas pela hidroxilamina? Por que sim ou por que não?

26. A sequência de nucleotídios a seguir é encontrada em um segmento curto de DNA:

5′-ATGT-3′
3′-TACA-5′

Se esta sequência for tratada com hidroxilamina, quais sequências surgiriam após a replicação?

*27. A sequência de nucleotídios a seguir é encontrada em um segmento curto de DNA:

5′-AG-3′
3′-TC-5′

a. Indique todas as sequências mutantes que podem surgir a partir da depurinação espontânea nesse segmento de DNA.
b. Indique todas as sequências mutantes que podem surgir a partir da desaminação espontânea nesse segmento de DNA.

28. Em muitos organismos eucarióticos, uma proporção significativa de bases citosina é naturalmente metilada para 5-metilcitosina. Com a evolução, a proporção de pares de base AT no DNA nesses organismos cresceu. Você pode sugerir algum possível mecanismo para esse aumento?

Seção 18.3

*29. Um químico sintetizou quatro novos compostos químicos no laboratório e os denominou PFI1, PFI2, PFI3 e PFI4. Ele fornece os compostos PFI a uma geneticista amiga e pede que ela determine seu potencial mutagênico. A geneticista descobre que os quatro compostos são extremamente mutagênicos. Ela também testa a capacidade de as mutações provocadas pelos compostos PFI serem revertidas por outros mutágenos conhecidos e obtém os resultados a seguir. Quais conclusões você pode tirar sobre a natureza das mutações produzidas por esses compostos?

	Revertidas por			
Mutações provocadas por	**2-aminopurina**	**Ácido nitroso**	**Hidroxilamina**	**Acridina laranja**
PFI1	Sim	Sim	Algumas	Não
PFI2	Não	Não	Não	Não
PFI3	Sim	Sim	Não	Não
PFI4	Não	Não	Não	Sim

ANÁLISE DE DADOS

30. Mary Alexander estudou os feitos da radiação sobre as taxas de mutação nos espermatozoides da *Drosophila melanogaster*. Ela irradiou as larvas da *Drosophila* com 3.000 roentgens (r) ou 3.975 r, coletou os machos adultos que se desenvolveram a partir das larvas irradiadas, cruzou-os com fêmeas não irradiadas e então contou o número de moscas mutantes F_1 produzidas por cada macho. Todas as moscas mutantes que surgiram foram usadas em subsequentes cruzamentos para determinar se seus fenótipos mutantes eram genéticos. A pesquisadora obteve os seguintes resultados (M. L. Alexander. 1954. *Genetics* 39:409-428):

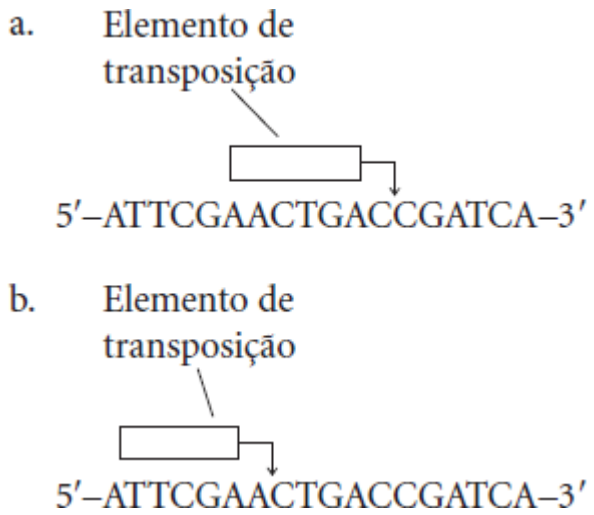
Grupo	**Número de descendentes**	**Descendentes com uma mutação genética**
Controle (0 r)	45.504	0
Irradiados (3.000 r)	49.512	71
Irradiados (3.975 r)	50.159	70

a. Calcule as taxas de mutação do grupo-controle e dos dois grupos de moscas irradiadas.
b. Com base nesses dados, você acha que a radiação influencia a mutação? Explique sua resposta.

31. Que conclusões você tiraria se o número de colônias de bactérias na Figura 18.21 fosse o mesmo que na placa-controle e na placa de tratamento? Explique seu raciocínio.
32. Um preceptor de genética projeta um experimento laboratorial para estudar os efeitos da radiação UV sobre a mutação nas bactérias. No experimento, os estudantes expõem as bactérias semeadas em placas de Petri à luz UV por diferentes períodos de tempo, colocam as placas em uma incubadora por 48 h e depois contam o número de colônias em cada placa. As placas que receberam mais radiação UV devem ter mais dímeros de pirimidina, que bloqueiam a replicação; assim, menos colônias devem surgir nas placas expostas à luz UV por períodos de tempo mais longos. Antes da realização do experimento, o preceptor avisa que enquanto as bactérias estiverem na incubadora, é essencial que os estudantes não abram a porta da incubadora exceto se a sala estiver escura. Por que as bactérias não devem ser expostas à luz?

Seção 18.4

*33. Determinado elemento de transposição gera repetições diretas flanqueadoras que têm 4 pb de comprimento. Indique a sequência que será encontrada em ambos os lados do elemento de transposição se esse elemento se inserir na posição indicada em cada uma das seguintes sequências.



34. Os olhos brancos na *Drosophila melanogaster* resultam de uma mutação recessiva ligada ao X. Em alguns casos, os mutantes de olhos brancos dão origem a descendentes de olhos brancos com pequenas manchas vermelhas. O número, a distribuição e o tamanho das manchas vermelhas são variáveis. Explique como um elemento de transposição poderia ser responsável por esse fenômeno.
35. Qual fator poderia determinar o comprimento das repetições diretas flanqueadoras que são produzidas na transposição?
*36. Quais dos pares de sequência a seguir podem ser encontrados nas extremidades de uma sequência de inserção?
    a. 5′-GGGCCAATT-3′ e 5′-CCCGGTTAA-3′
    b. 5′-AAACCCTTT-3′ e 5′-AAAGGGTTT-3′
    c. 5′-TTTCGAC-3′ e 5′-CAGCTTT-3′
    d. 5′-ACGTACG-3′ e 5′-CGTACGT-3′
    e. 5′-GCCCCAT-3′ e 5′-GCCCAT-3′
37. Explique por que a semente do milho na Figura 18.31 D é variegada, com algumas áreas coloridas e outras áreas sem pigmento.
*38. Duas cepas diferentes de *Drosophila melanogaster* são submetidas a cruzamentos recíprocos. Quando os machos da cepa A são cruzados com as fêmeas da cepa B, os descendentes são normais. Entretanto, quando as fêmeas da cepa A são cruzadas com machos da cepa B, existem muitas mutações e rearranjos cromossômicos nos gametas dos descendentes F_1 e a geração F_1 é estéril. Explique esses resultados.
39. Uma sequência de inserção tem uma grande deleção em seu gene transposase. Em quais circunstâncias essa sequência de inserção conseguiria transpor?
40. A zidovudina (AZT) é um fármaco usado para tratar pacientes com AIDS. A AZT bloqueia a enzima transcriptase reversa usada pelo vírus da imunodeficiência humana (HIV), o agente responsável pela AIDS. Você espera que AZT tenha algum efeito nos elementos de transposição? Se sim, que tipo de elementos de transposição seriam afetados e qual seria o efeito mais provável?
41. Foi encontrado um elemento de transposição que codifica uma enzima transcriptase reversa. Com base nessa informação, que conclusões você pode tirar sobre a provável estrutura e o método de transposição desse elemento?

*42. Um geneticista examina uma espiga de milho na qual a maioria dos grãos são amarelos, mas ele encontra alguns com manchas púrpura, como apresentado aqui. Forneça uma possível explicação para o surgimento das manchas púrpura nesses grãos de milho, de acordo com seus diferentes tamanhos. (Dica: consulte a Seção sobre os elementos *Ac* e *Ds* no milho.)

Seção 18.5

43. Qual mecanismo de reparo do DNA seria mais provável de corrigir o erro de incorporação marcado pelo balão 2 na Figura 18.10?

*44. Um agricultor quer isolar mutantes nos tomates que apresentam defeito no reparo do DNA. Entretanto, ele não tem conhecimento nem equipamentos para estudar as enzimas nos sistemas de reparo do DNA. Como ele pode identificar as plantas que têm deficiência no reparo do DNA? Quais são os traços que ele deve procurar?

Questões desafiadoras

Seção 18.1

ANÁLISE DE DADOS

45. Robert Bost e Richard Cribbs estudaram uma cepa de *E. coli* (*araB14*) que tinha uma mutação *nonsense* no gene estrutural que codifica a L-ribuloquinase, uma enzima que possibilita que a bactéria metabolize o açúcar arabinose (R. Bost and R. Cribbs. 1969. *Genetics* 62:1-8). A partir da cepa *araB14*, eles isolaram algumas bactérias que tinham mutações que faziam com que elas voltassem para o tipo selvagem. A análise genética desses revertentes mostrou que eles tinham duas mutações supressoras diferentes. Uma mutação supressora (*R1*) estava ligada à mutação original na L-ribuloquinase e provavelmente ocorria no mesmo *locus*. Por si só, essa mutação possibilitou a produção de L-ribuloquinase, mas a enzima não era tão efetiva para metabolizar arabinose como a enzima codificada pelo alelo do tipo selvagem. A segunda mutação supressora (Su^B) não estava ligada à mutação original. Junto com a mutação *R1*, Su^B tornou possível a produção de L-ribuloquinase, mas Su^B por si só não era capaz de suprimir a mutação original.
    a. Com base nessas informações, as mutações *R1* e Su^B eram supressores intragênicos ou intergênicos? Justifique sua resposta.
    b. Proponha uma explicação para como *R1* e Su^B restauram a capacidade de *araB14* metabolizar arabinose e por que Su^B é capaz de restaurar completamente essa capacidade.

46. Acondroplasia é um distúrbio autossômico dominante caracterizado por baixa estatura desproporcional: os braços e as pernas são curtos em comparação com a cabeça e o tronco. Esse distúrbio é causado por uma substituição de base no gene, localizado no braço curto do cromossomo 4, para o receptor 3 do fator de crescimento de fibroblasto (FGFR3). Embora esteja claro que a acondroplasia é herdada como um traço autossômico dominante, mais de 80% das pessoas com essa condição têm genitores com estatura normal. Essa alta porcentagem indica que a maioria dos casos é causada por mutações recém-surgidas; esses casos (não herdados de um genitor afetado) são chamados de esporádicos. Estudos demonstraram que os casos esporádicos de acondroplasia são quase sempre causados por mutações herdadas do pai (mutações paternas). Além disso, a ocorrência de acondroplasia é maior entre os pais mais velhos; aproximadamente 50% das crianças com acondroplasia nascem de pais com mais de 35 anos. Não existe associação com a idade da mãe. A taxa de mutação para acondroplasia (cerca de 4×10^{-5} mutações por gameta) é elevada em comparação com a taxa de outros distúrbios genéticos. Explique por que a maioria das mutações espontâneas para acondroplasia tem origem no pai e por que a ocorrência é maior em homens mais velhos.

47. A doença de Tay-Sachs é uma doença genética autossômica recessiva grave, que provoca surdez, cegueira, convulsões e morte. Essa doença resulta de um defeito no gene *HEXA*, que codifica hexosaminidase A. Essa enzima normalmente degrada os gangliosídios G_{M2}. Quando não existe hexosaminidase A, os gangliosídios G_{M2} se acumulam no cérebro. Os resultados de estudos moleculares recentes mostram que a mutação mais comum que provoca essa doença é uma inserção de 4 pb que produz um códon de parada prematuro *downstream*. Os resultados de outros estudos revelaram que a transcrição do gene *HEXA* é normal em pessoas com essa doença, mas o mRNA do *HEXA* é instável. Proponha um mecanismo para explicar como um códon de parada prematuro poderia causar instabilidade do mRNA.

48. *Ocre* e *âmbar* são dois tipos de mutações *nonsense*. Antes de o código genético ser desvendado, Sydney Brenner, Anthony O. Stretton e Samuel Kaplan aplicaram diferentes tipos de mutágenos aos bacteriófagos em uma tentativa para determinar as bases nos códons responsáveis pelas mutações *âmbar* e *ocre*. Eles sabiam que essas mutações eram suprimidas por diferentes tipos de mutações, demonstrando que cada uma é um códon de término diferente. Eles obtiveram os seguintes resultados:
    1. A substituição de base única poderia converter uma mutação *ocre* em uma mutação *âmbar*.
    2. A hidroxilamina induzia mutações *ocre* e *âmbar* nos fagos do tipo selvagem.
    3. 2-aminopurina fez *ocre* mudar para *âmbar*.
    4. Hidroxilamina não fez *ocre* mudar para *âmbar*.

    Esses dados não possibilitam a determinação da sequência completa de nucleotídios dos códons *âmbar* e *ocre*, mas fornecem informações sobre as bases encontradas nas mutações *nonsense*.
    a. Que conclusões podem ser tiradas sobre as bases encontradas nos códons das mutações âmbar e ocre a partir dessas observações?
    b. Dos três códons *nonsense* (UAA, UAG, UGA), qual representa a mutação *ocre*?

Seção 18.3

49. Para determinar se a radiação associada às bombas atômicas de Hiroshima e Nagasaki provocou mutações recessivas na linhagem germinativa, os cientistas examinaram a razão sexual dos descendentes dos sobreviventes dos ataques. Você pode explicar por que seria esperado que mutações germinativas alterassem a razão sexual?

Seção 18.4

50. Marilyn Houck e Margaret Kidwell propuseram que os elementos P eram carreados da *Drosophila willistoni* para a *Drosophila melanogaster* por ácaros que se alimentam dessas moscas (M. A. Houck *et al*. 1991. *Science* 253:1125-1129). Que evidência você acha que seria necessária para demonstrar que a *D. melanogaster* adquiriu os elementos *P* dessa maneira? Proponha vários experimentos para provar tal evidência.

Seção 18.5

51. A tricotiodistrofia é um distúrbio humano hereditário caracterizado por envelhecimento prematuro, incluindo osteoporose, osteosclerose, cabelos grisalhos prematuros, infertilidade e redução da expectativa de vida. Os resultados dos estudos mostraram que a mutação que causa esse distúrbio ocorre em um gene que codifica uma DNA helicase. Proponha um mecanismo sobre como uma mutação na DNA helicase poderia provocar envelhecimento prematuro. Certifique-se de relacionar as manifestações do distúrbio com possíveis funções da enzima helicase.

19 Análise da Genética Molecular e Biotecnologia

Como ajudar o cego a enxergar

A engenharia genética, a manipulação e transferência de genes de um organismo para outro, tem revolucionado o estudo da genética e conseguido criar uma vasta gama de novos produtos, incluindo bactérias que geram substâncias químicas, plantações resistentes a pragas e animais de fazenda que secretam produtos farmacêuticos no seu leite. Em sua principal aplicação, a engenharia genética está sendo usada para tratar doenças nos humanos, um processo conhecido como terapia gênica. Em um exemplo dramático de terapia gênica, em 2007, pesquisadores transferiram genes para quatro pessoas cegas, restaurando parcialmente sua visão.

A visão é um dos sentidos mais preciosos, a qual nos possibilita ler, contornar obstáculos físicos, reconhecer os amigos e desfrutar a estonteante beleza visual do mundo natural. Infelizmente, milhões de pessoas são cegas ou têm limitações visuais. A visão frequentemente deteriora com a idade, e a maioria dos casos de cegueira ocorre em idosos. Entretanto, algumas crianças nascem cegas, e outras perdem a visão cedo. A pesquisa sugere que a hereditariedade é responsável por cerca de 50% dos casos de cegueira antes dos 45 anos. Por causa da complexidade do olho, de seus nervos associados e das partes do cérebro responsáveis pela percepção visual, defeitos em vários genes podem levar à cegueira.

A terapia gênica tem sido aplicada a uma forma genética rara de cegueira conhecida como amaurose congênita de Leber (ACL), na qual há uma mutação em um dos vários genes responsáveis pelo desenvolvimento ou função dos receptores de luz na retina. A ACL é herdada como um distúrbio autossômico recessivo; nas pessoas com duas cópias defeituosas de um gene, as células sensíveis à luz morrem. As crianças com ACL começam a perder a visão ao nascer e a maioria está completamente cega aos 40 anos.

Um tipo de ACL é causado por um defeito no gene *RPE65*, que codifica uma enzima que ajuda a converter a vitamina A em rodopsina, um pigmento que absorve luz. Sem a enzima, a rodopsina não é produzida e as células fotorreceptoras atrofiam com o tempo. Em 1998, os pesquisadores descobriram que alguns cães da raça Swedish Briard também têm um defeito no *RPE65* e ficam cegos. A pesquisa nestes cães possibilitou melhor compreensão de como os defeitos no *RPE65* provocam cegueira e sugeriu uma possível solução, a terapia gênica. Em 2001, os pesquisadores trataram três cães jovens com a terapia gênica. A visão dos cães melhorou o suficiente para que eles pudessem evitar objetos e andassem em um labirinto.

Em 2007, os pesquisadores na Pensilvânia e Londres usaram a terapia gênica para tratar quatro adultos jovens com ACL. Cada paciente recebeu uma injeção de vírus geneticamente modificado que carreava uma cópia funcional de *RPE65*. Os resultados foram incríveis: todos os pacientes mostraram melhora significativa na percepção visual. Alguns que antes eram capazes de detectar apenas o movimento das mãos puderam ler várias linhas em uma carta de olho. Um paciente que não conseguia desvencilhar-se dos obstáculos em seu trajeto, passou a fazê-lo após o tratamento. Todos os quatro pacientes ainda eram legalmente cegos, mas os resultados indicavam que a terapia gênica pode ser efetiva para ACL. Os pesquisadores agora usaram a terapia em outras pessoas com ACL; os resultados foram impressionantes, sobretudo nos pacientes jovens, alguns dos quais a visão melhorou o suficiente para que eles pudessem praticar esportes.

Este capítulo introduz algumas das técnicas usadas na engenharia genética. Vamos começar analisando a tecnologia da genética molecular e alguns de seus efeitos. Então vamos examinar vários métodos usados para isolar, estudar, alterar e recombinar sequências de DNA e recolocá-las nas células. Finalmente, vamos explorar algumas das aplicações da análise da genética molecular.

19.1 As técnicas de genética molecular revolucionaram a biologia

Em 1973, um grupo de cientistas produziu os primeiros organismos com moléculas de DNA recombinante. Stanley Cohen na Stanford University e Herbert Boyer na University of California School of Medicine em São Francisco e seus colegas inseriram

um fragmento de DNA de um plasmídio em outro, criando uma molécula de DNA totalmente nova, recombinante. Então eles introduziram o plasmídio recombinante em *E. coli*. Esses experimentos lideraram uma das revoluções mais importantes na história da ciência.

A **tecnologia do DNA recombinante** é um conjunto de técnicas moleculares para localizar, isolar, alterar e estudar segmentos de DNA. O termo *recombinante* é usado porque o objetivo é combinar o DNA de duas fontes distintas. Os genes de duas bactérias diferentes podem ser unidos, por exemplo, ou um gene humano pode ser inserido em um cromossomo viral. Comumente chamada de **engenharia genética**, a tecnologia do DNA recombinante agora inclui muitas técnicas moleculares que podem ser usadas para analisar, alterar e recombinar praticamente qualquer sequência de DNA de várias fontes.

Revolução da genética molecular

As técnicas da tecnologia do DNA recombinante são apenas parte de uma vasta gama de métodos moleculares atualmente disponíveis. Estas técnicas moleculares alteraram drasticamente a forma como os genes são estudados. Antes, as informações sobre a estrutura e a organização dos genes eram obtidas pelo exame de seus efeitos fenotípicos, mas a análise da genética molecular permite que as próprias sequências de nucleotídios sejam lidas. Esta análise fornece novas informações sobre a estrutura e a função dos genes e alterou muitos conceitos fundamentais da genética. Nossa compreensão detalhada dos processos genéticos, como replicação, transcrição, tradução, processamento do DNA e regulação gênica foi obtida com o uso de técnicas de genética molecular. Estas técnicas são usadas em muitos outros campos também, incluindo bioquímica, microbiologia, biologia de desenvolvimento, neurobiologia, evolução e ecologia.

A tecnologia do DNA recombinante e outras técnicas moleculares também estão sendo usadas para criar vários produtos comerciais, incluindo fármacos, hormônios, enzimas e plantações. Uma indústria inteira, a **biotecnologia**, surgiu com o uso destas técnicas para desenvolver novos produtos. Na Medicina, a genética molecular é usada para sondar a natureza do câncer, diagnosticar doenças genéticas e infecciosas, produzir fármacos e tratar distúrbios hereditários.

Conceitos

A genética molecular e a tecnologia do DNA recombinante são usadas para localizar, analisar, alterar, estudar e recombinar sequências de DNA. Essas técnicas são usadas para investigar a estrutura e a função dos genes, abordar questões em muitas áreas da biologia, criar produtos comerciais e diagnosticar e tratar doenças.

Como trabalhar no nível molecular

A manipulação dos genes no nível molecular é um grande desafio, muitas vezes exigindo estratégias que, no primeiro momento, não parecem óbvias. O problema básico é que os genes são muito pequenos, e cada célula tem milhares deles. Os nucleotídios não podem ser vistos individualmente, e nenhuma característica física marca o início ou o fim de um gene.

Analisemos uma situação típica enfrentada pelo geneticista molecular. Suponha que queiramos usar bactérias para produzir grandes quantidades de uma proteína humana. O primeiro e mais formidável problema é encontrar o gene que codifica a proteína desejada. Um genoma diploide humano tem 3,2 bilhões de pares de base no DNA. Vamos considerar que o gene que queremos isolar tem 3.000 pb de comprimento. Nosso gene-alvo ocupa apenas um milionésimo do genoma, então pesquisá-lo no imenso espaço do DNA genômico é mais difícil que encontrar uma agulha no palheiro. Mas, mesmo se formos capazes de localizar o gene, como vamos separá-lo do resto do DNA?

Se conseguirmos localizar e isolar o gene desejado, precisaríamos inseri-lo em uma célula bacteriana. Fragmentos lineares de DNA são resumidamente degradados pelas bactérias, então o gene deve ser inserido em uma forma estável. Ele também deve ser capaz de se replicar ou não será transmitido na divisão celular. Se conseguirmos transferir nosso gene para as bactérias em uma forma estável, precisamos ter certeza de que ele será adequadamente transcrito e traduzido.

Finalmente, os métodos usados para isolar e transferir genes são ineficientes e, de um milhão de células sujeitas a esses procedimentos, apenas *uma* célula poderia conseguir captar e expressar o gene humano. Então é preciso pesquisar em muitas células bacterianas para descobrir uma com o DNA recombinante. E voltamos para o problema da agulha em um palheiro.

Embora esses problemas pareçam insuperáveis, foram desenvolvidas técnicas moleculares para contorná-los. Genes humanos são rotineiramente transferidos para as células bacterianas, onde são expressos.

Conceitos

As análises de genética molecular demandam métodos especiais porque os genes individuais compõem uma fração mínima do DNA celular e não podem ser vistos.

19.2 Técnicas moleculares são usadas para isolar, recombinar e ampliar os genes

A primeira etapa na análise molecular de um segmento de DNA ou gene é isolá-lo do resto do DNA e fazer muitas cópias de modo que possam ser feitas mais análises. O isolamento e a ampliação do DNA exigem que ele seja combinado com outras moléculas de DNA. Nas seções a seguir, vamos examinar algumas das técnicas moleculares que usamos para isolar, recombinar e amplificar segmentos de DNA.

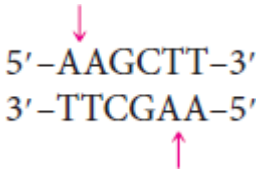
Corte e união dos fragmentos de DNA

Um evento importante no desenvolvimento dos métodos de genética molecular foi a descoberta no final dos anos 1960 das **enzimas de restrição** (também chamadas de **endonucleases de restrição**) que reconhecem e fazem cortes de fita dupla no DNA em sequências de nucleotídios específicas. Essas enzimas são produzidas naturalmente pelas bactérias e usadas na defesa contra vírus. Uma bactéria protege seu próprio DNA de uma enzima de restrição ao modificar a sequência de reconhecimento, em geral ao adicionar grupos metila a seu DNA.

Já foram isolados vários tipos diferentes de enzimas de restrição das bactérias. As enzimas de restrição do tipo II reconhecem sequências específicas e cortam o DNA em sítios definidos na sequência de reconhecimento ou perto dela. Praticamente todo trabalho de genética molecular é feito com enzimas de restrição do tipo II; as discussões sobre as enzimas de restrição neste livro referem-se às enzimas do tipo II.

Já foram isolados das bactérias mais de 800 tipos diferentes de enzimas que reconhecem e cortam o DNA em mais de 100 diferentes sequências. Muitas destas enzimas estão disponíveis no comércio; o **Quadro 19.1** apresenta exemplos de algumas enzimas de restrição comumente usadas. O nome de cada enzima de restrição começa com uma abreviação que significa sua origem bacteriana.

As sequências reconhecidas pelas enzimas de restrição têm 4 a 8 pb de comprimento, a maioria das enzimas reconhece uma sequência de 4 a 6 pb. A maioria das sequências de reconhecimento é palindrômica – sequências que leem o mesmo (de 5′ para 3′) nas duas fitas DNA complementares.



Algumas das enzimas fazem cortes desencontrados no DNA. Por exemplo, *Hind*III reconhece a seguinte sequência:

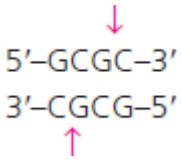
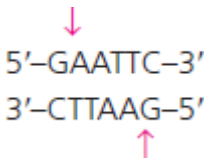
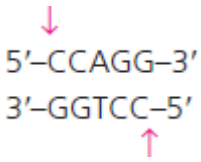
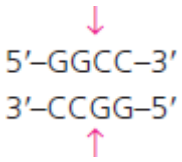
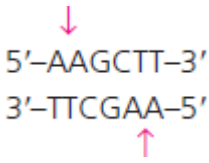
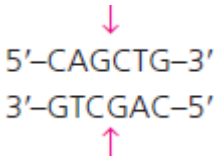
*Hind*III corta o esqueleto açúcar-fosfato de cada fita no ponto indicado pela seta, gerando fragmentos com extremidades protuberantes, curtas, de fita única:

5′-A AGCTT-3′
3′-TTGGA A-5′

Tais extremidades são chamadas de **extremidades coesivas** ou extremidades adesivas porque elas são complementares a cada uma e podem parear espontaneamente para ligar os fragmentos. Assim, os fragmentos de DNA podem ser “colados”: dois fragmentos quaisquer rompidos pela mesma enzima terão extremidades complementares e vão parear (**Figura 19.1**). Quando suas extremidades coesivas pareiam, dois fragmentos de DNA podem ser unidos de forma permanente pela DNA ligase, que sela os cortes entre os grupos de açúcar-fosfato dos fragmentos.

Quadro 19.1 Características de algumas enzimas de restrição do tipo II usadas na tecnologia do DNA recombinante.

Enzima	Microrganismo que produz a enzima	Sequência de reconhecimento	Tipo de extremidade de fragmento produzida
*Bam*HI	*Bacillus amyloliquefaciens*	5′–GGATCC–3′ 3′–CCTAGG–3′	Coesiva

*Cof*I	*Clostridium formicoaceticum*		Coesiva
*Eco*RI	*Escherichia coli*		Coesiva
*Eco*RII	*Escherichia coli*		Coesiva
*Hae*III	*Haemophilus aegyptius*		Cega
*Hin*dIII	*Haemophilus influenzae*		Coesiva
*Pvu*II	*Proteus vulgaris*		Cega

Nota: As três primeiras letras da abreviação para cada enzima de restrição referem-se à espécie de bactéria da qual a enzima foi isolada (p. ex., Eco refere-se a *E. coli*). Uma quarta letra pode se referir à cepa da bactéria da qual a enzima foi isolada (o “R” no *Eco*RI significa que esta enzima foi isolada da cepa RY13 da *E. coli*). Os algarismos romanos que acompanham as letras identificam enzimas diferentes da mesma espécie.

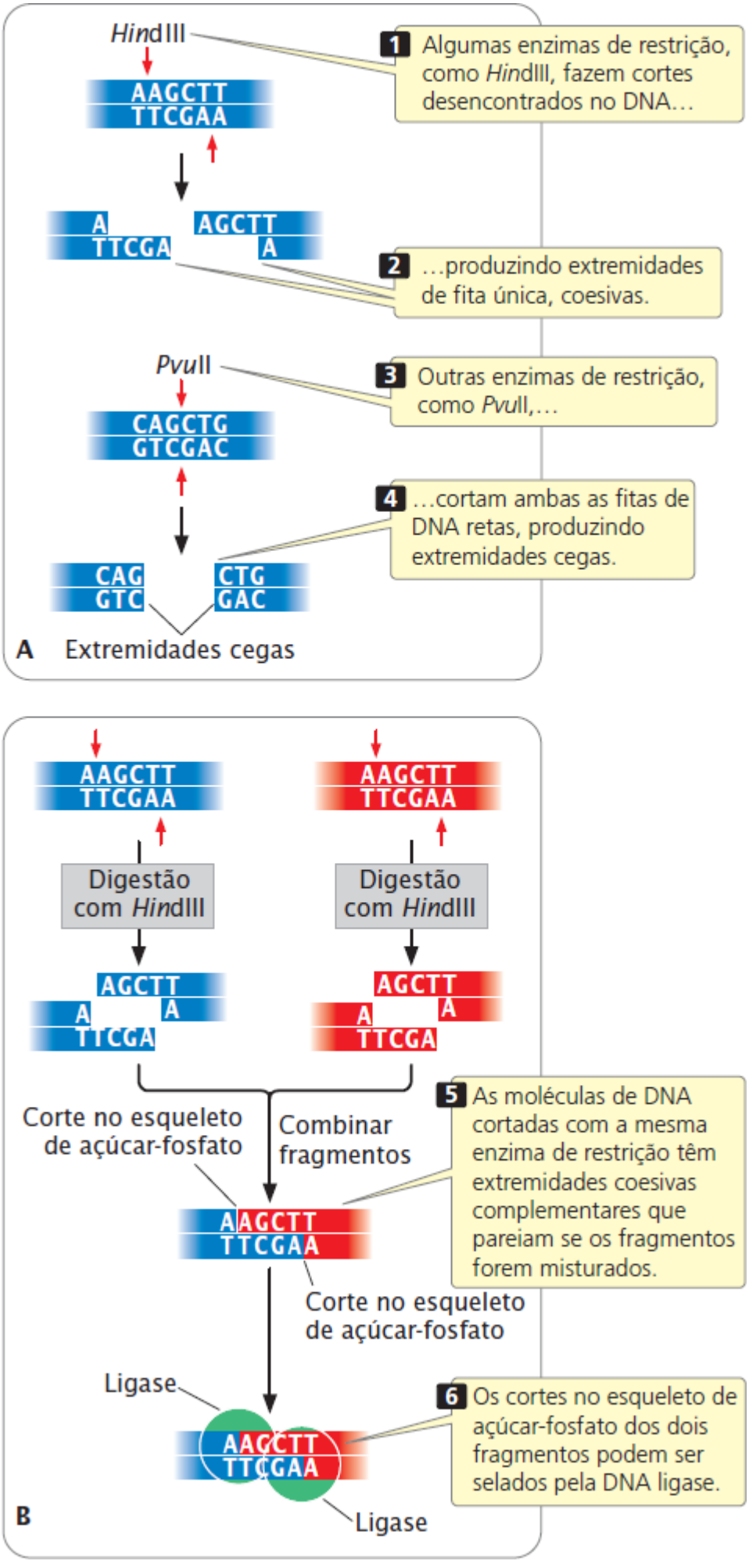
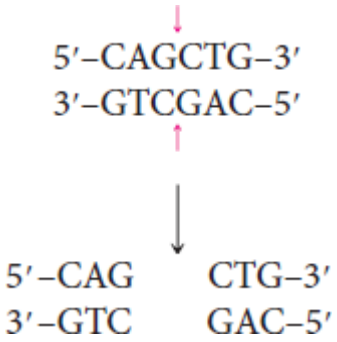


Figura 19.1 As enzimas de restrição fazem cortes de fita dupla no DNA, produzindo extremidades coesivas.

Nem todas as enzimas de restrição produzem cortes desencontrados e extremidades coesivas. *Pvu*II corta no meio do sítio de reconhecimento e os cortes nas duas fitas são diretamente opostas uma à outra, produzindo fragmentos de extremidades cegas:



Os fragmentos com extremidades cegas são unidos de várias formas.

As sequências reconhecidas por uma enzima de restrição estão localizadas aleatoriamente no genoma. Assim, existe uma relação entre o comprimento da sequência de reconhecimento e o número de vezes que ela está presente em um genoma: existem menos sequências de reconhecimento maiores do que as sequências de reconhecimento menores porque a probabilidade de ocorrência de uma determinada sequência composta, por exemplo, de seis bases específicas é menor que a probabilidade de ocorrência de uma sequência de quatro bases específicas. Consequentemente, as enzimas de restrição que reconhecem sequências mais longas cortarão um fragmento específico de DNA em poucos fragmentos mais longos que as enzimas de restrição que reconhecem sequências mais curtas.

As enzimas de restrição são usadas sempre que os fragmentos de DNA precisam ser cortados ou unidos. Em uma típica reação de restrição, uma solução concentrada de DNA purificado é colocada em um pequeno tubo com uma solução tampão e uma pequena quantidade de enzima de restrição. A mistura reacional é então aquecida na temperatura adequada para a enzima, em geral 37°C. Após algumas horas, a enzima corta os sítios de restrição adequados em todas as moléculas de DNA, produzindo um conjunto de fragmentos de DNA.

Nos últimos anos, os geneticistas desenvolveram enzimas mais complexas, chamadas de **nucleases modificadas**, que produzem cortes de fita dupla no DNA em qualquer sequência de DNA predeterminada. As nucleases modificadas são parte de uma enzima de restrição que cliva o DNA, acoplada com outra proteína que reconhece e se liga a uma sequência específica de DNA; essa sequência específica é determinada pela sequência de aminoácidos da proteína. Ao alterar a sequência de aminoácidos da proteína ligadora, a nuclease modificada pode ser personalizada ao se ligar e cortar qualquer sequência específica de DNA. As nucleases modificadas incluem as nucleases dedo de zinco (ZFNs), que usam uma proteína ligadora de DNA chamada dedo de zinco, e as nucleases ativadoras de transcrição (TALENs), que usam uma proteína que normalmente se liga a sequências nos promotores. Resolva o Problema 29

Conceitos

As enzimas de restrição tipo II cortam o DNA em sequências de bases específicas palindrômicas. Algumas enzimas de restrição fazem cortes desencontrados, produzindo fragmentos de DNA com extremidades coesivas, outras cortam ambas as fitas, produzindo fragmentos de extremidades cegas. Existem menos sequências de reconhecimento longas no DNA do que sequências de reconhecimento curtas.

Checagem dos conceitos 1

Qual a origem das enzimas de restrição?

Visualização dos fragmentos de DNA

Após o término de uma reação de restrição, surgem várias dúvidas. A enzima de restrição cortou o DNA? O DNA foi cortado em quantos fragmentos? Quais são os tamanhos dos fragmentos produzidos? A eletroforese em gel fornece meios para responder a estas perguntas.

A eletroforese é uma técnica bioquímica padrão para separar as moléculas com base no seu tamanho e carga elétrica. Existem vários tipos diferentes de eletroforese; a **eletroforese em gel** é usada para separar as moléculas de DNA. Um gel poroso é preparado a partir de agarose (um polissacarídio isolado das algas), derretido em uma solução tampão e transferido para um molde de plástico. À medida que ele esfria, a agarose solidifica, fazendo com que o gel lembre uma gelatina dura.

São feitos pequenos poços em uma extremidade do gel para manter as soluções dos fragmentos de DNA (**Figura 19.2 A**) e

uma corrente elétrica passa pelo gel. Como o grupo fosfato de cada molécula de DNA tem carga elétrica negativa, os fragmentos de DNA migram para a extremidade positiva do gel. Nesta migração, o gel poroso funciona como uma peneira, separando os fragmentos de DNA por tamanho. Os pequenos fragmentos de DNA migram mais rapidamente do que os maiores, e com o passar do tempo, os fragmentos se separam com base no seu tamanho. Os fragmentos de DNA de tamanho conhecido (padrões de tamanhos) são colocados em outro poço. Ao comparar a distância de migração de fragmentos desconhecidos com a distância viajada pelos padrões de tamanho, um pesquisador pode determinar o tamanho aproximado dos fragmentos desconhecidos (**Figura 19.2 B**).

Os fragmentos de DNA ainda são muitos pequenos para serem visualizados, então é necessário resolver a questão de visualizar o DNA. É possível visualizar de várias formas. O procedimento mais simples é corar o gel com um corante específico para os ácidos nucleicos, como o brometo de etídio, que penetra firmemente (intercala) entre as bases de DNA e fluoresce quando exposto à luz UV, produzindo bandas brilhantes no gel (**Figura 19.2 C**). Por outro lado, os fragmentos de DNA podem ser visualizados ao adicionar um marcador ao DNA antes de colocá-los no gel. Por exemplo, marcadores químicos podem ser detectados ao adicionar anticorpos ou outras substâncias que carreiam um corante e se fixarão ao DNA em questão, que pode ser visualizado diretamente. **Resolva o Problema 30**

Conceitos

Os fragmentos de DNA podem ser separados e seus tamanhos podem ser determinados com o uso de eletroforese em gel. Os fragmentos podem ser visualizados ao usar um corante específico para ácidos nucleicos ou marcar os fragmentos com um corante químico.

Checagem dos conceitos 2

Os fragmentos de DNA com 500, 1.000 e 2.000 pb de comprimento são separados por eletroforese em gel. Qual fragmento migrará mais rápido pelo gel?

a. Fragmento de 2.000 pb.
b. Fragmento de 1.000 pb.
c. Fragmento de 500 pb.
d. Todos vão migrar a mesma distância.

Localização dos fragmentos de DNA com *Southern blotting* e sondas

Se um fragmento pequeno de DNA, como um plasmídio, é cortado por uma enzima de restrição, os poucos fragmentos produzidos podem ser observados como bandas distintas em um gel de eletroforese. Ao contrário, se o DNA genômico de uma célula é cortado por uma enzima de restrição, são produzidos vários fragmentos de diferentes tamanhos. Uma enzima de restrição que reconhece uma sequência de quatro bases cortaria teoricamente a cada 256 pb. O genoma humano, com 3,2 bilhões de pares de bases, geraria mais de 12 milhões de fragmentos quando cortado por esta enzima de restrição. Separado por eletroforese e corado, este grande conjunto de fragmentos apareceria como um esfregaço contínuo no gel por causa da presença de tantos fragmentos de diferentes tamanhos. Em geral, os pesquisadores estão interessados em apenas alguns desses fragmentos, talvez os carreadores de um gene específico. Como eles podem localizar os fragmentos desejados nesta imensa mistura de DNA?

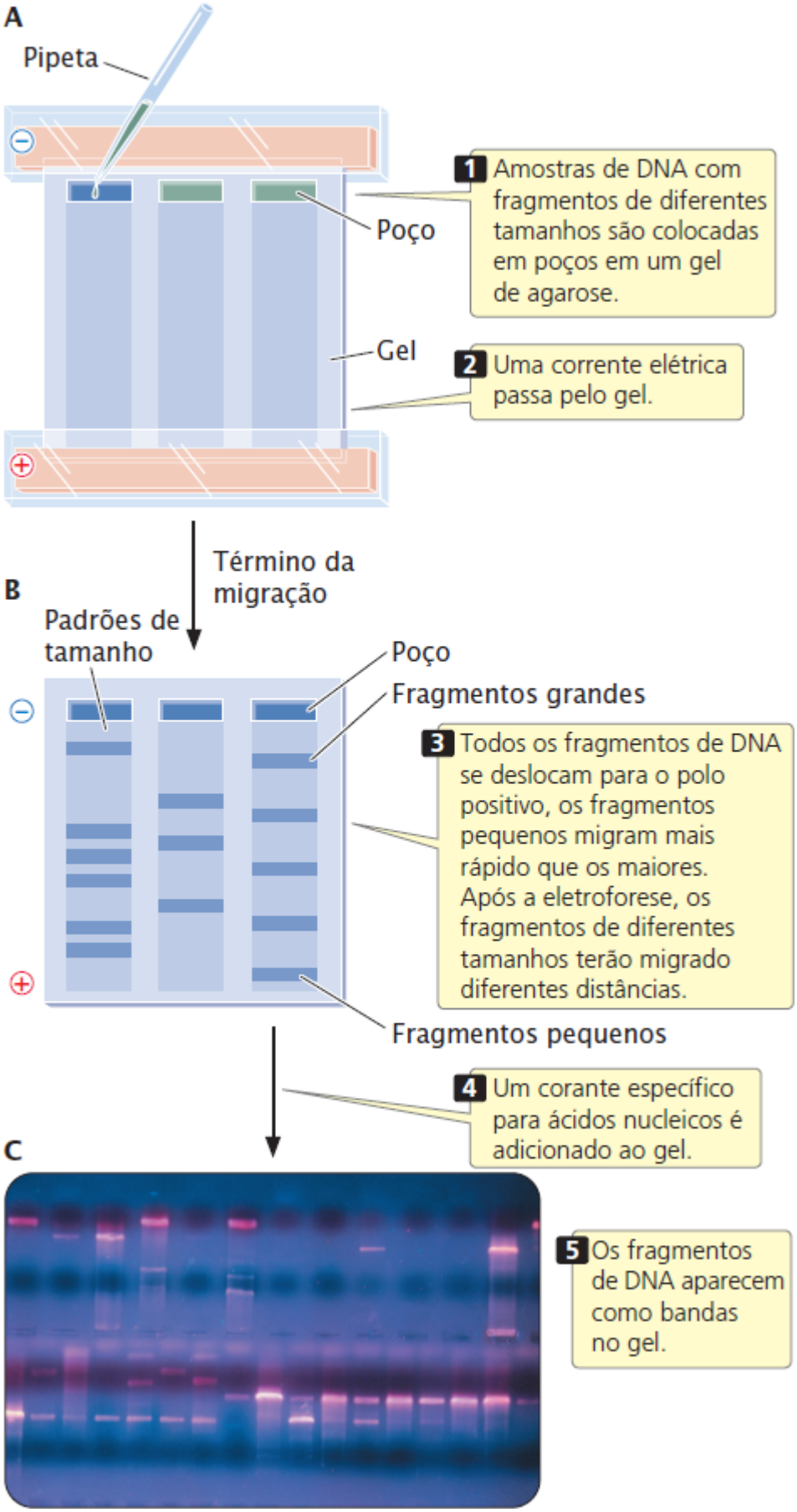
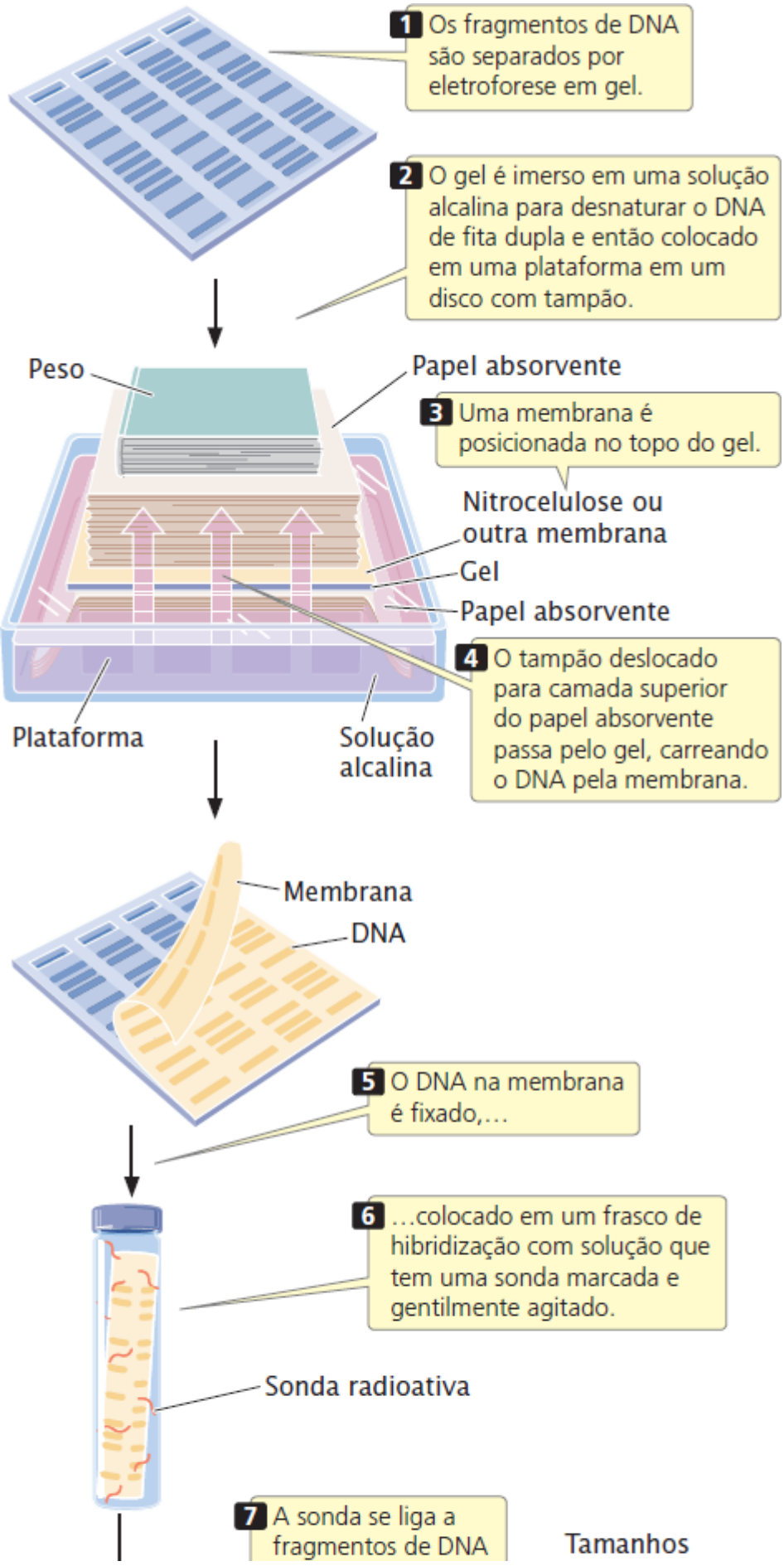


Figura 19.2 A eletroforese em gel pode ser usada para separar a molécula de DNA com base no seu tamanho e na carga elétrica. (Fotografia: Klaus Guldbrandsen/Science Source.)

Uma possibilidade é usar uma **sonda**, uma molécula de DNA ou RNA com uma sequência de base complementar a uma sequência no gene de interesse. As bases em uma sonda pareiam apenas com as bases em uma sequência complementar e, se adequadamente marcada, a sonda pode ser usada para localizar um gene específico ou outra sequência de DNA. Para usar uma sonda, primeiro o pesquisador corta o DNA em fragmentos ao usar uma ou mais enzimas de restrição e então separa os

fragmentos com eletroforese em gel (**Figura 19.3**). A seguir, os fragmentos separados serão desnaturados e transferidos para um meio sólido permanente (como nitrocelulose ou membrana de náilon). O ***Southern blotting*** (nome dado em homenagem a Edwin M. Southern) é uma técnica usada para transferir os fragmentos de fita única, desnaturados, de um gel para um meio sólido permanente.



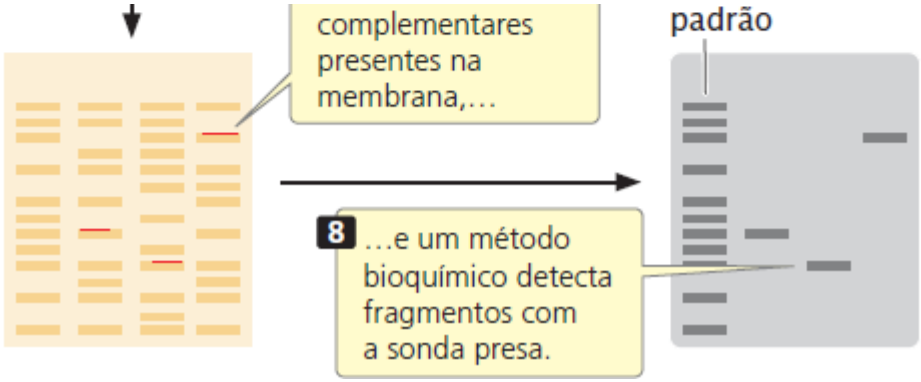


Figura 19.3 O *Southern blotting* e a hibridização com sondas podem localizar alguns fragmentos específicos em um grande conjunto de DNA.

Após os fragmentos de DNA de fita única serem transferidos, a membrana é colocada em uma solução de hibridização de uma sonda marcada (ver Figura 19.3). A sonda se ligará a qualquer fragmento de DNA na membrana que tenha as sequências complementares. Uma sonda se liga apenas a uma parte do fragmento de DNA e então este fragmento pode ter sequências não encontradas na sonda. A membrana é lavada para remover qualquer sonda não ligada, e um método bioquímico revela a presença de sonda ligada.

O RNA pode ser transferido de um gel para um suporte sólido usando um procedimento semelhante chamado ***Northern blotting*** (não é homenagem, mas o termo é usado para combinar com Southern). A hibridização de uma sonda pode revelar o tamanho de determinada molécula de mRNA, sua abundância relativa ou os tecidos nos quais o mRNA é transcrito. ***Western blotting*** é a transferência de proteína de um gel para uma membrana. Aqui, a sonda é um anticorpo, usado para determinar o tamanho de uma proteína específica e o padrão de expressão da proteína.

Conceitos

As sondas marcadas, sequências de RNA ou DNA complementares à sequência de interesse, podem ser usadas para localizar genes individuais ou sequências de DNA. O *Southern blotting* pode ser usado para transferir fragmentos de DNA de um gel para uma membrana como nitrocelulose.

Checagem dos conceitos 3

Qual a diferença do *Northern* e *Western blotting* para o *Southern blotting*?

Clonagem de genes

Muitos métodos de DNA recombinante precisam de várias cópias de um fragmento específico de DNA. Uma forma para ampliar um fragmento específico de DNA é colocar o fragmento em uma célula bacteriana e deixar que a célula replique o DNA. Este procedimento é chamado de **clonagem de genes** porque são produzidas cópias idênticas (clones) do fragmento original de DNA.

Um **vetor de clonagem** é uma molécula de DNA estável, replicante, à qual um fragmento de DNA exógeno foi preso para introduzir em uma célula. Um vetor de clonagem efetivo tem três características importantes (**Figura 19.4**): (1) uma origem de replicação, que garante que o vetor seja replicado dentro da célula; (2) marcadores selecionáveis, que permitem que qualquer célula com o vetor seja selecionada ou identificada e (3) um ou mais sítios de restrição em que um fragmento de DNA possa ser inserido. Os sítios de restrição usados para clonagem devem ser únicos; se um vetor é cortado em múltiplos sítios de reconhecimento, são gerados vários fragmentos de DNA e é possível organizá-los na ordem correta, mas é muito difícil.

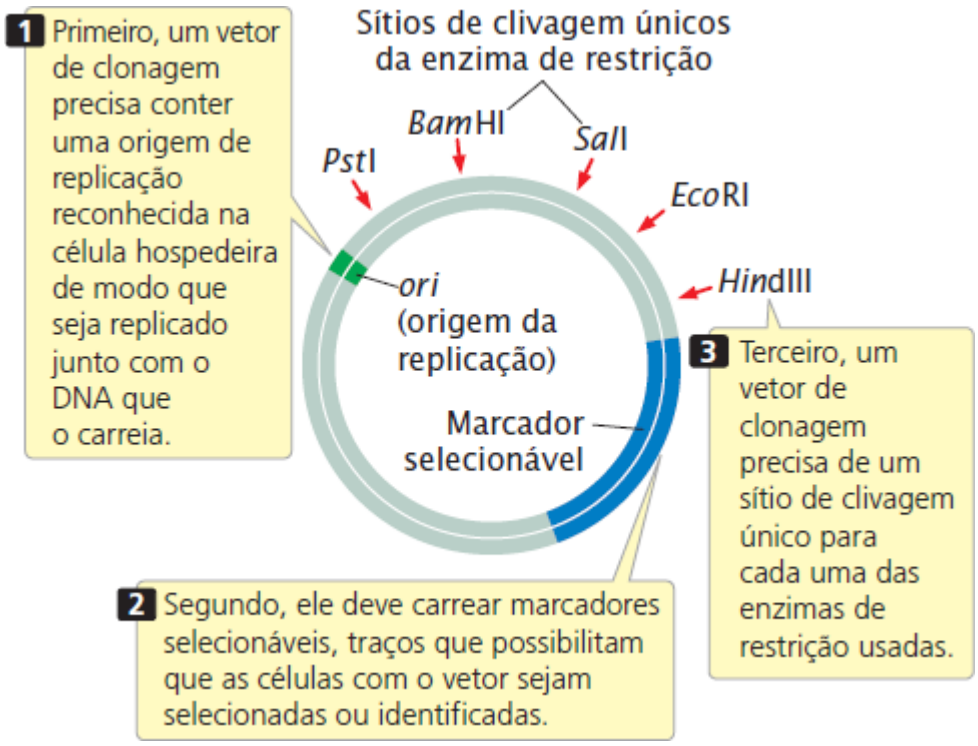


Figura 19.4 Um vetor de clonagem idealizado tem uma origem de replicação, um ou mais marcadores selecionáveis e sítios para uma ou mais enzimas de restrição.

Vetores plasmídios. Os plasmídios, moléculas de DNA circulares que existem naturalmente nas bactérias (ver *Capítulo 9*), são usados como vetores para clonar fragmentos de DNA nas bactérias. Eles têm origens de replicação e são capazes de se replicar independentemente do cromossomo bacteriano. Os plasmídios usados na clonagem são construídos a partir de plasmídios bacterianos naturais, maiores e têm múltiplos sítios de restrição, uma origem do sítio de replicação e marcadores selecionáveis (**Figura 19.5**).

O método mais fácil para inserir um gene em um vetor plasmídio é cortar o DNA estranho (com o gene) e o plasmídio com a mesma enzima de restrição (**Figura 19.6**). Se a enzima de restrição fizer cortes desencontrados no DNA, são produzidas extremidades coesivas complementares no DNA exógeno e no plasmídio. O DNA e o plasmídio são misturados, e parte dos fragmentos de DNA exógeno vai parear com as extremidades cortadas do plasmídio. A DNA ligase é usada para selar os cortes no esqueleto de açúcar-fosfato, criando um plasmídio recombinante que tem o fragmento de DNA exógeno.

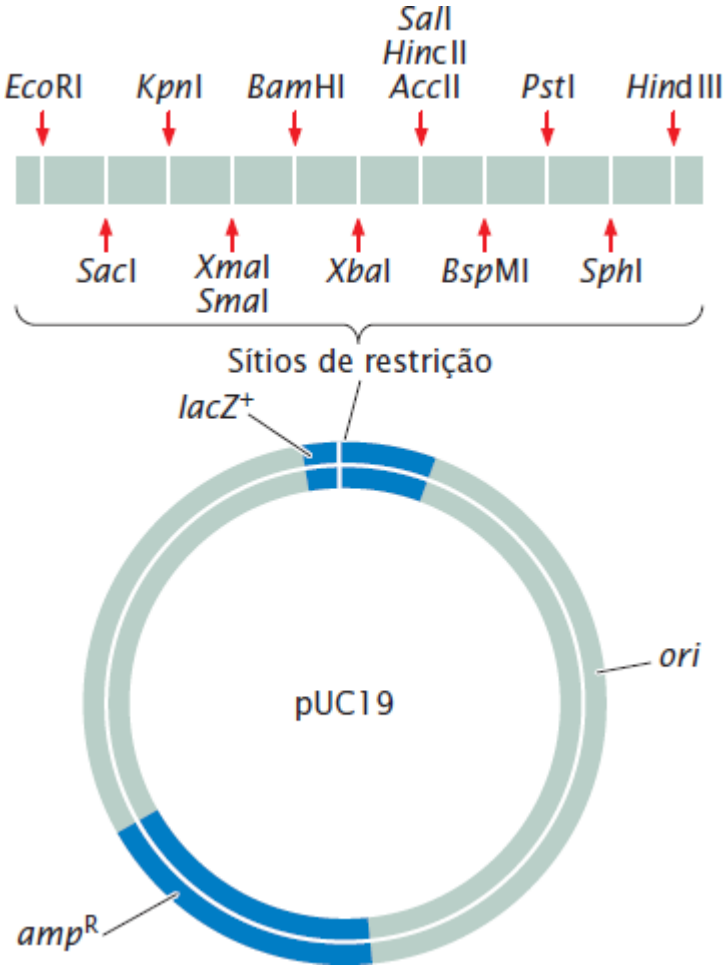


Figura 19.5 O plasmídio pUC19 é um típico vetor de clonagem. Ele tem um grupo de sítios de restrição único, uma origem de replicação e dois marcadores selecionáveis – um gene que confere resistência a ampicilina e um gene *lacZ*.

Às vezes não há sítios de restrição em um sítio onde o DNA precisa ser cortado. Nesse caso, um sítio de restrição pode ser criado com o uso de sítios de clonagem ou **ligadores (*linkers*)**, que são pequenos fragmentos sintéticos de DNA com um ou mais sítios de restrição. Os *linkers* podem ser fixados às extremidades de qualquer fragmento de DNA e então são cortados por uma enzima de restrição, gerando extremidades coesivas complementares às extremidades coesivas no plasmídio.

Transformação. Quando um gene precisa ser colocado dentro de um plasmídio, o plasmídio deve ser introduzido nas células bacterianas. Esta tarefa é feita por *transformação*, na qual as células bacterianas captam o DNA do ambiente externo (ver *Capítulo 9*). Alguns tipos de células sofrem transformação naturalmente; outros sofrem tratamento químico ou físico antes de serem transformados. Dentro da célula, os plasmídios se replicam e multiplicam e as próprias células se multiplicam.

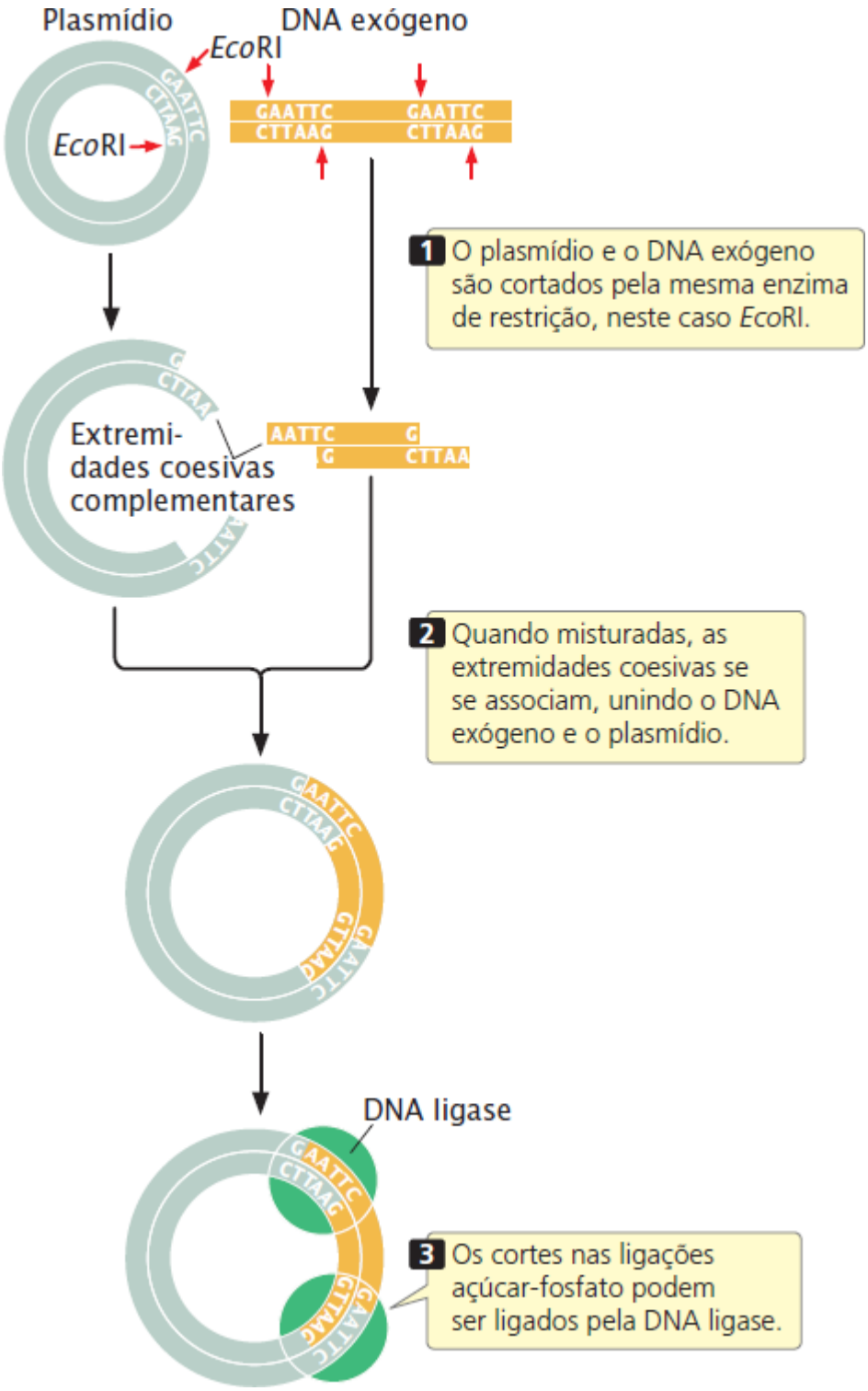


Figura 19.6 **Um fragmento exógeno de DNA pode ser inserido em um plasmídio com o uso de enzimas de restrição.**

Seleção de células para plasmídios recombinantes. As células que carreiam os plasmídios recombinantes podem ser detectadas com o uso de marcadores selecionáveis no plasmídio. Os marcadores selecionáveis comuns são genes que conferem resistência a um antibiótico: qualquer célula que tenha tal plasmídio será capaz de viver na presença do antibiótico, que normalmente mataria as células bacterianas.

Uma forma comum de rastrear células à procura de um plasmídio recombinante é usar um plasmídio que tenha um fragmento do gene *lacZ*, uma pequena parte da extremidade inicial do gene (**Figura 19.7**). Esse gene *lacZ* parcial tem vários sítios de restrição singulares nos quais um fragmento de DNA pode ser inserido e clonado. As bactérias que fazem o trabalho de clonagem têm características especiais que evidenciam o plasmídio recombinante. As bactérias que serão transformadas pelo plasmídio são *lacZ*$^{-}$, sem a extremidade inicial do gene *lacZ*, mas com sua extremidade terminal. Sem o plasmídio, as bactérias são incapazes de sintetizar β-galactosidase. Se nenhum DNA exógeno for inserido no gene *lacZ* parcial do plasmídio, a extremidade inicial do gene (fornecida pelo plasmídio) e a extremidade terminal do gene (fornecida pelas bactérias) trabalham juntas dentro da célula bacteriana para produzir a β-galactosidase. Se o DNA exógeno for inserido de forma eficiente no sítio de restrição, ele rompe a extremidade inicial do gene *lacZ* e a β-galactosidase não é produzida. O plasmídio também tem um

marcador selecionável, que pode ser um gene que confere resistência a um antibiótico como ampicilina.

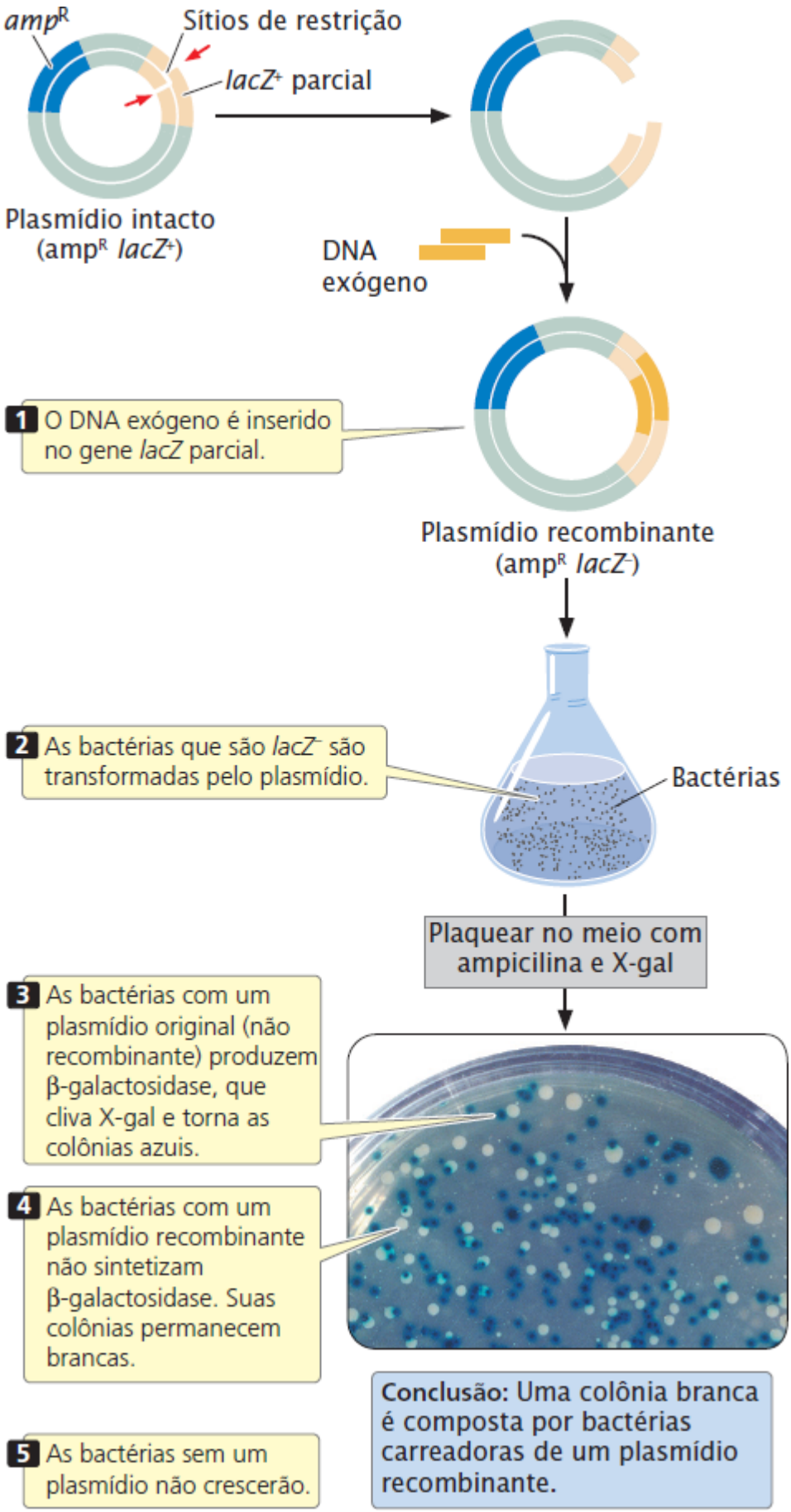


Figura 19.7 O gene *lacZ* pode ser usado para rastrear bactérias contendo plasmídios recombinantes. Um plasmídio especial carreia um fragmento do gene lacZ e um gene que confere resistência à ampicilina. (Fotografia cortesia de Edvotek. www.edvotek.com)

As bactérias *lacZ*⁻ são transformadas pelos plasmídios e plaqueadas em um meio com ampicilina. Apenas as células que

foram transformadas de fato e tenham um plasmídio com o gene que confere resistência à ampicilina sobrevivem e crescem. Algumas destas células terão um plasmídio intacto, enquanto outras terão um plasmídio recombinante. O meio também tem a substância química X-gal, que produz uma substância azul quando é clivada. As células com um plasmídio original intacto, sem um fragmento inserido, têm um gene *lacZ* funcional (a extremidade inicial do gene fornecida pelo plasmídio e a extremidade terminal fornecida pelas bactérias). Estas bactérias podem sintetizar β-galactosidase, que cliva X-gal e torna as bactérias azuis. As células bacterianas com um plasmídio recombinante, entretanto, têm uma extremidade inicial do gene da β-galactosidase rompido pelo DNA inserido; elas não sintetizam esta enzima e permanecem brancas. (Nestes experimentos, o próprio gene da β-galactosidase da bactéria está inativo, e apenas as bactérias com o plasmídio ficam azuis.) Assim, a cor da colônia da bactéria permite a rápida determinação se um plasmídio recombinante ou intacto estiver presente na célula. Após as células com o plasmídio recombinante serem identificadas, elas podem crescer em grande escala, replicando o fragmento inserido do DNA.

Os plasmídios produzem vetores de clonagem ideais, mas podem reter apenas DNA de menos de 15 kb de tamanho. O vetor plasmídio fica instável quando são inseridos grandes fragmentos de DNA.

Outros vetores de genes. Foram desenvolvidos vários outros vetores para clonar fragmentos grandes de DNA nas bactérias. Por exemplo, o bacteriófago λ, que infecta a *E. coli*, pode ser usado para clonar até 23.000 pb de DNA exógeno: ele transfere o DNA para as células bacterianas com alta eficiência. Os **cosmídios** são plasmídios empacotados em revestimentos proteicos virais vazios e transferidos para bactérias por meio da infecção viral. Eles podem carrear mais de duas vezes o DNA estranho que um vetor fago. Os **cromossomos bacterianos artificiais** (BACs) são vetores originalmente construídos a partir do plasmídio F (um plasmídio especial que controla o acasalamento e transferência de material genético em algumas bactérias, ver *Capítulo 9*) e podem reter fragmentos muito grandes de DNA de até 300.000 pb. O **Quadro 19.2** compara as propriedades dos plasmídios, vetores de fago, cosmídios e BACs.

Em alguns casos, o objetivo da clonagem de genes não é apenas replicar o gene, mas também produzir a proteína que ele codifica. Para garantir a transcrição e a tradução, um gene exógeno é inserido em um **vetor de expressão**, que, além da origem comum de replicação, sítios de restrição e marcadores selecionáveis, tem sequências necessárias para transcrição e tradução nas células bacterianas (**Figura 19.8**).

Quadro 19.2 Comparação dos plasmídios, vetores do fago lambda, cosmídios e cromossomos bacterianos artificiais.

Vetor de clonagem	Tamanho do DNA que pode ser clonado	Métodos de propagação	Introdução a bactérias
Plasmídio	Até 15 kb	Replicação de plasmídio	Transformação
Fago lambda	Até 23 kb	Reprodução do fago	Infecção pelo fago
Cosmídio	Até 44 kb	Reprodução do plasmídio	Infecção pelo fago
Cromossomo bacteriano artificial	Até 300 kb	Reprodução do plasmídio	Eletroporação

Nota: 1 kb = 1.000 pb. Eletroporação são pulsos elétricos que aumentam a permeabilidade de uma membrana.

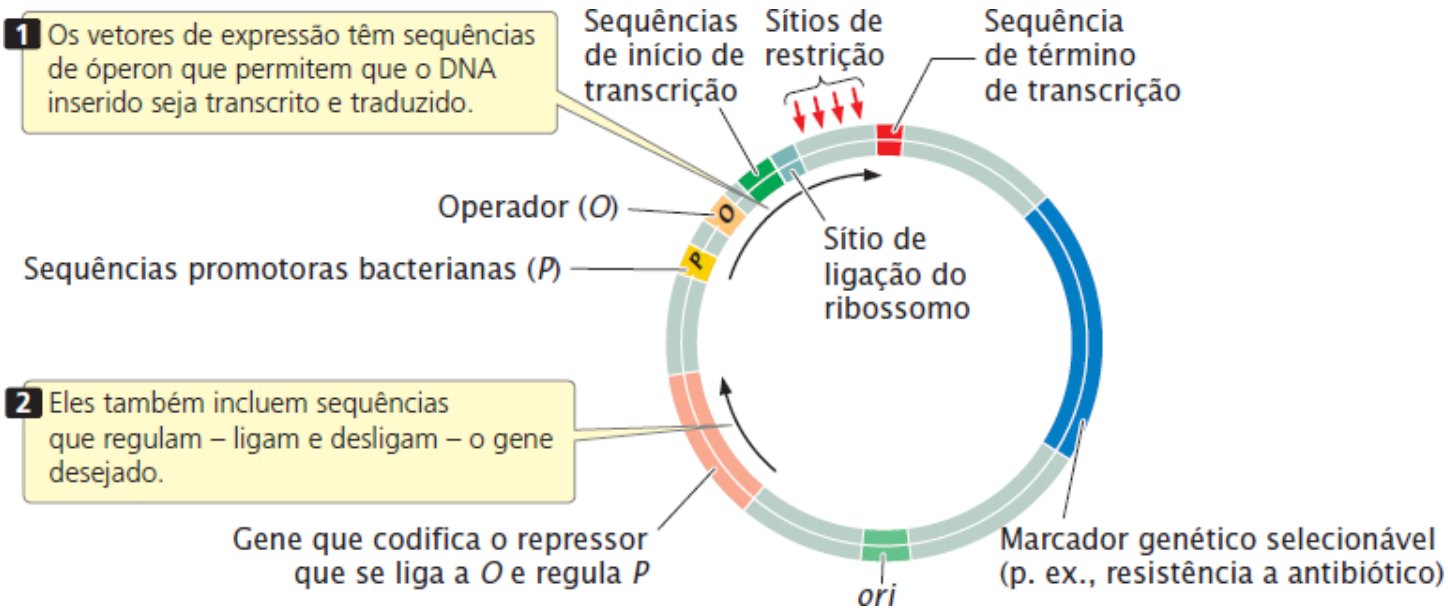


Figura 19.8 Para garantir a transcrição e tradução, um gene exógeno pode ser inserido em um vetor de expressão, neste exemplo, um vetor de expressão da *E. coli*.

Embora a manipulação de genes nas bactérias seja simples e eficiente, o objetivo pode ser transferir um gene para as células eucarióticas. Por exemplo, pode-se desejar transferir um gene que confere resistência a herbicida a uma plantação ou transferir um gene para o fator de coagulação de uma pessoa que sofra de hemofilia. Muitas proteínas eucarióticas são modificadas após a tradução (p. ex., podem ser adicionados grupos açúcar). Tais modificações são essenciais para o funcionamento correto, mas as bactérias não conseguem executar a modificação, então uma proteína funcional só pode ser produzida em uma célula eucariótica.

Foram desenvolvidos vários vetores de clonagem que permitem a inserção dos genes nas células eucarióticas. Foram desenvolvidos plasmídios especiais para clonagem na levedura e vetores retrovirais para clonagem nos mamíferos.

Um **cromossomo artificial de levedura** (YAC) é uma molécula de DNA com uma origem de replicação, um par de telômeros e um centrômero. Estas características garantem que os YACs sejam estáveis, se repliquem e se separem da mesma forma que os cromossomos de leveduras. Eles são muito úteis porque podem carrear fragmentos de DNA até 600 kb e alguns YACs especiais podem carrear inserções de mais de 1.000 kb. Os YACS são modificados de modo que podem ser usados em organismos eucarióticos além da levedura.

A bactéria de solo *Agrobacterium tumefaciens*, que invade as plantas através das feridas e induz a galha da coroa (tumores), tem sido usada para transferir genes para plantas. Esta bactéria tem um grande plasmídio chamado **plasmídio Ti**, parte do qual é transferido para uma célula da planta quando *A. tumefaciens* a infecta. Na planta, parte do DNA do plasmídio Ti se integra a um dos cromossomos da planta, onde é transcrito e traduzido para produzir várias enzimas que ajudam a manter a bactéria (**Figura 19.9 A**).

Os geneticistas desenvolveram um vetor com sequências flanqueadoras necessárias para transferir o DNA (chamadas de TL e TR, ver Figura 19.9 A), um marcador selecionável e sítios de restrição nos quais o DNA exógeno pode ser inserido (**Figura 19.9 B**). Quando colocado em uma *A. tumefaciens* com um plasmídio Ti auxiliar (que carreia as sequências que permitem que a bactéria infecte a célula da planta, o que o vetor artificial não tem), o vetor transferirá o DNA exógeno que carreia para uma célula da planta, onde se integrará em um cromossomo da planta. Este vetor tem sido usado para transferir genes que conferem características economicamente importantes, como resistência a herbicidas, vírus de plantas e pragas. Resolva o Problema 31

Conceitos

Os fragmentos de DNA podem ser inseridos em vetores de clonagem, pequenos fragmentos de DNA que se replicam dentro de uma célula. Um vetor de clonagem deve ter uma origem de replicação, um ou mais sítios de restrição únicos e marcadores selecionáveis. Um vetor de expressão tem sequências que permitem que um gene clonado seja transcrito e traduzido. Foram desenvolvidos vetores de clonagem especiais para introduzir os genes nas células eucarióticas.

Checagem dos conceitos 4

Como um gene é inserido em um vetor plasmídio de clonagem?

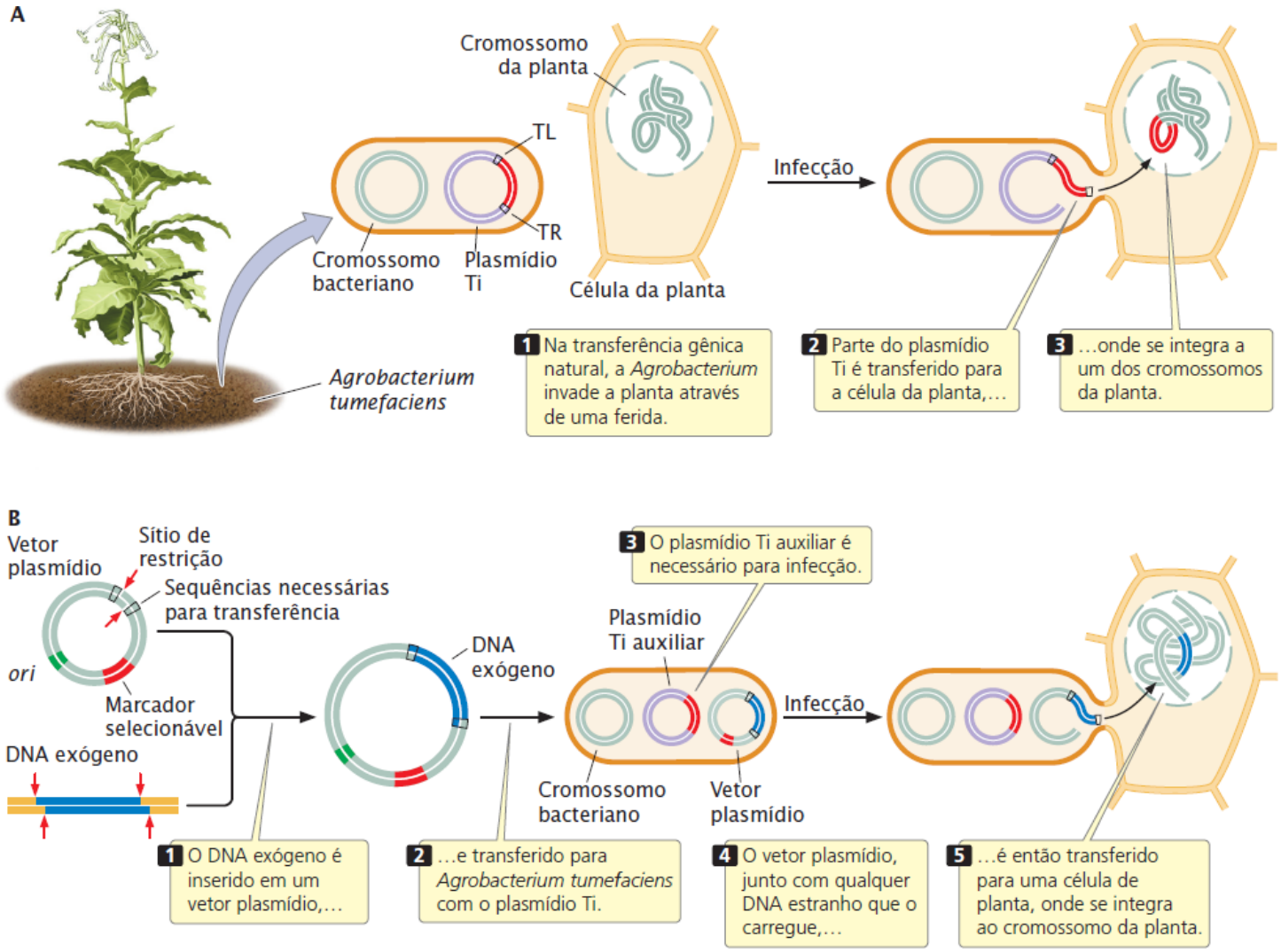


Figura 19.9 O plasmídio Ti pode ser usado para transferir genes para as plantas. As sequências flanqueadoras TL e TR são necessárias para a transferência de segmento de DNA da bactéria para a célula da planta.

Aplicação | Engenharia genética das plantas com pesticidas

Um uso pioneiro e economicamente importante das técnicas moleculares para transferência de genes a outros organismos é a engenharia genética de plantas que produzem seus próprios inseticidas. A bactéria *Bacillus thuringiensis* produz naturalmente uma proteína, conhecida como toxina Bt, letal para muitos insetos. Essa toxina é particularmente atraente como inseticida porque é específica para alguns insetos, degradada rapidamente no ambiente e não tóxica para os humanos e outros animais.

Em 1987, Mark Vaeck e seus colaboradores no *Plant Genetic Systems* na Bélgica fizeram manipulação genética nas plantas do tabaco para expressar a toxina Bt. Outros pesquisadores isolaram o gene que codifica a toxina Bt da *B. thuringiensis* e o clonaram em um plasmídio da *E. coli* de forma que as sequências genéticas poderiam ser manipuladas com facilidade (**Figura 19.10**). Vaeck e seus colaboradores usaram enzimas de restrição para romper as sequências do gene *Bt* destes plasmídios em fragmentos de diferentes tamanhos. Os fragmentos do gene *Bt* foram conectados a um gene (*neo*) que confere resistência ao antibiótico canamicina, que é tóxico para as plantas e outros eucariotos, fornecendo um marcador selecionável para as células com o gene *Bt*. Essas sequências sintéticas (chamadas de construtos gênicos) têm fragmentos do gene *Bt* de diferentes comprimentos ligados a um gene *neo* (ver Figura 19.10). Os construtos foram inseridos em um vetor de expressão que tinha um promotor, para garantir que as sequências introduzidas seriam transcritas. O vetor de expressão também tinha sequências consenso poli(A), que garantiam que o mRNA produzido a partir dos genes *Bt* e *neo* unidos seriam processados adequadamente e transcritos nas células das plantas.

As bactérias *Agrobacterium tumefaciens* foram então transformadas com os vetores de expressão. Após os vetores penetrarem em *Agrobacterium*, as sequências no vetor se recombinaram com as sequências em um plasmídio Ti, transferindo os construtos gênicos para este plasmídio. Pequenos discos de folhas de uma planta do tabaco foram infectados com a *Agrobacterium* geneticamente modificada, que transferiu os plasmídios Ti para as células da planta. Plantas inteiras de tabaco foram regeneradas a partir dos discos de folhas e selecionadas em função da resistência à canamicina.

As plantas transgênicas resultantes foram então testadas quanto à resistência a pragas de insetos. As folhas das plantas foram

usadas para alimentar lagarta do tabaco e as taxas de mortalidade foram monitoradas. Foi observada alta mortalidade das lagartas em cerca de dois terços das plantas que tinham os fragmentos do gene *Bt*. Nenhuma das plantas com o gene *Bt* completo produzia qualquer toxina que mate insetos, aparentemente, as células eram melhores para traduzir os fragmentos do gene *Bt* do que traduzir o gene inteiro. As plantas transgênicas que produziam a toxina Bt cresciam normalmente. Elas foram intercruzadas para produzir as plantas F_1, que também exibiam resistência a insetos e canamicina, indicando que os genes introduzidos foram integrados de forma estável nos cromossomos da planta.

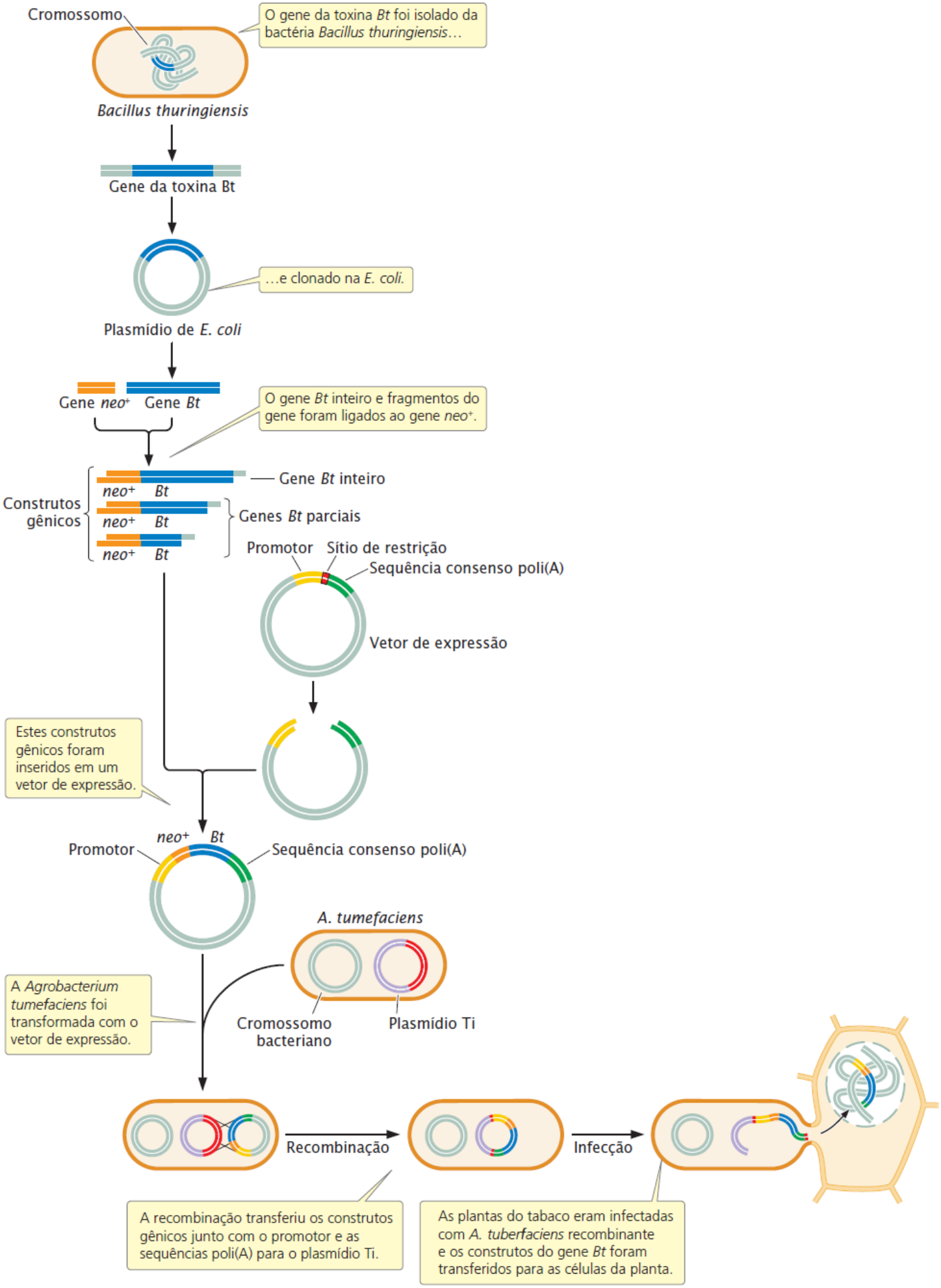


Figura 19.10 O gene para toxina *Bt*, que codifica um inseticida, foi isolado a partir das bactérias e transferido para as plantas do tabaco.

Após a pesquisa feita por Vaeck e seus colaboradores, outros pesquisadores usaram métodos semelhantes para introduzir o gene para a toxina Bt no algodão, tomate, milho e outras plantas. Atualmente, as plantas transgênicas que expressam a toxina Bt são cultivadas em vários países.

Amplificação dos fragmentos de DNA com reação em cadeia da polimerase

A manipulação e análise dos genes exigem múltiplas cópias das sequências de DNA usadas. Uma questão crítica em trabalhar em nível molecular é que cada gene é uma fração mínima do DNA total da célula. Como cada gene é raro, ele deve ser isolado e amplificado antes que possa ser estudado. Uma forma para amplificar o fragmento de DNA é cloná-lo nas células bacterianas. De fato, por muitos anos, a clonagem de genes era a única forma de amplificar o fragmento de DNA e a clonagem era um pré-requisito para muitos outros métodos moleculares. Entretanto, a clonagem é um trabalho intenso e requer vários dias para a bactéria crescer. Atualmente, a reação em cadeia da polimerase possibilita a amplificação de fragmentos curtos de DNA sem clonagem, embora a clonagem ainda seja amplamente usada para amplificar fragmentos maiores de DNA e para outras manipulações de sequências de DNA.

A **reação em cadeia da polimerase** (PCR), desenvolvida pela primeira vez em 1983 por Kary Mullis, permite que os fragmentos de DNA sejam amplificados bilhões de vezes em algumas horas. Ela pode ser usada com quantidades extremamente pequenas do DNA original, mesmo uma única molécula. A reação em cadeia da polimerase revolucionou a biologia molecular e agora é uma das técnicas moleculares mais usadas. A base da PCR é a replicação catalisada por uma DNA polimerase. A replicação neste caso tem duas exigências essenciais: (1) um molde de DNA a partir do qual uma nova fita de DNA pode ser copiada e (2) um par de de iniciadores ou *primers* com um grupo 3′-OH ao qual podem ser adicionados novos nucleotídios.

Como uma molécula de DNA tem duas fitas de nucleotídios, cada uma pode servir como molde e a quantidade de DNA dobra em cada evento de replicação. Os *primers* usados na PCR para replicar os moldes são fragmentos curtos de DNA, em geral de 17 a 25 nucleotídios, que são complementares a sequências conhecidas no molde.

Reação em cadeia da polimerase. Para fazer PCR, os pesquisadores começam com uma solução que inclui o DNA-alvo, a DNA polimerase, todos os quatro trifosfatos de desoxirribonucleosídio (dNTPs, os substratos para a DNA polimerase), *primers* e íons magnésio e outros sais necessários para a reação. Uma típica reação em cadeia da polimerase inclui três etapas (**Figura 19.11**). Na etapa 1, uma solução inicial de DNA é aquecida entre 90° e 100°C para romper as pontes de hidrogênio entre as fitas e produz os moldes de fita única necessários. A mistura reacional é mantida nesta temperatura por apenas um minuto ou dois. Na etapa 2, a solução de DNA é resfriada rapidamente entre 30° e 65°C e mantida a esta temperatura por um minuto ou menos. Durante este curto intervalo, as fitas de DNA não têm chance de se reunir, mas os *primers* serão capazes de se fixar às fitas molde. Na etapa 3, a solução é aquecida por um minuto ou menos a 72°C, a temperatura na qual a DNA polimerase consegue sintetizar novas fitas de DNA. No final do ciclo, duas moléculas de DNA de fita dupla novas são produzidas para cada molécula original do DNA-alvo.

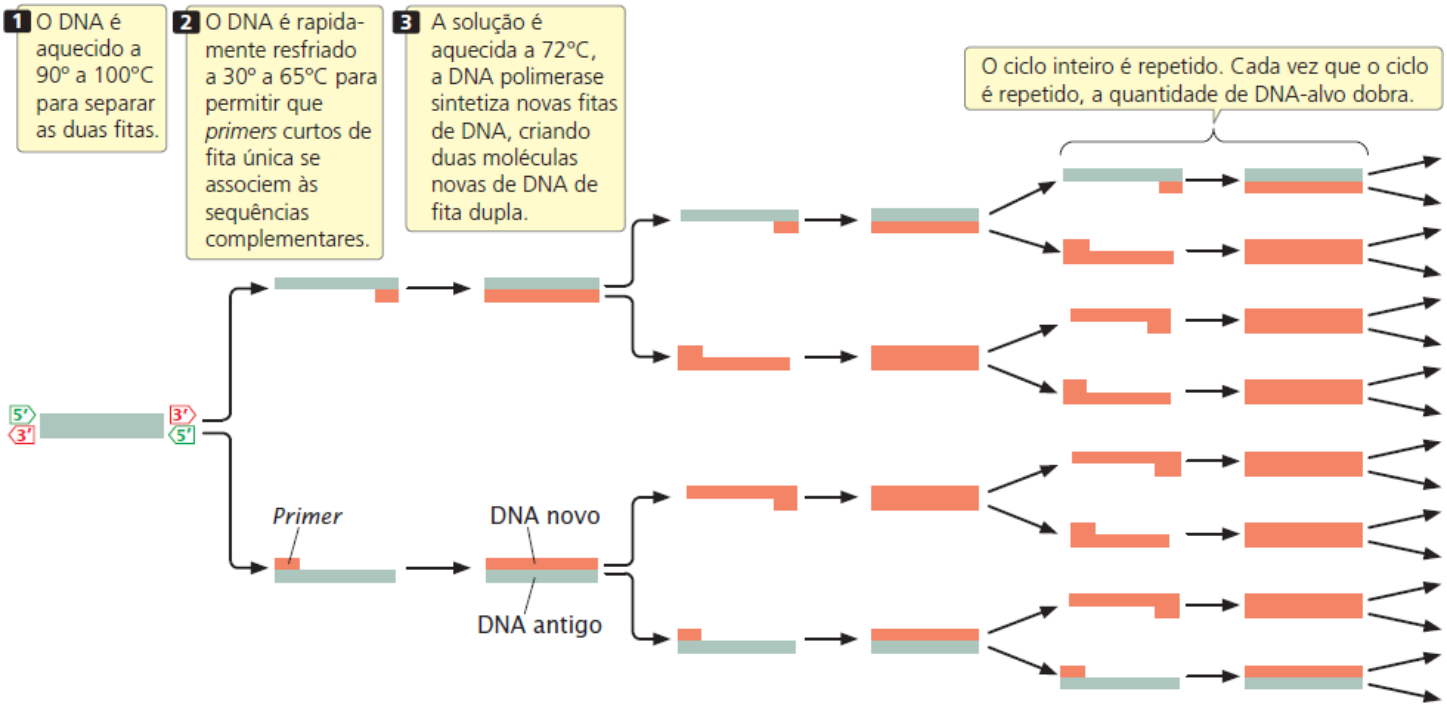


Figura 19.11 A reação em cadeia da polimerase é usada para amplificar até quantidades muito pequenas de DNA.

O ciclo inteiro é repetido. Com cada ciclo, a quantidade de DNA-alvo dobra, então a quantidade de DNA-alvo aumenta geometricamente. Uma molécula de DNA aumenta em mais de 1.000 moléculas em 10 ciclos de PCR, para mais de 1 milhão de moléculas em 20 ciclos e em mais de 1 bilhão de moléculas em 30 ciclos. Cada ciclo é completado dentro de alguns minutos, então pode ser obtida uma grande amplificação de DNA dentro de algumas horas.

Duas inovações fundamentais facilitaram o uso de PCR no laboratório. A primeira foi a descoberta de uma DNA polimerase que é estável a altas temperaturas usadas na etapa 1 da PCR. A DNA polimerase da *E. coli* que era usada na PCR desnatura a 90°C. Por este motivo, tinha de ser adicionada enzima nova à mistura reacional em *cada* ciclo, reduzindo consideravelmente a velocidade do processo. Este obstáculo foi superado quando a DNA polimerase foi isolada da bactéria *Thermus aquaticus*, que vive em águas termais no Parque Nacional de Yellowstone. Esta enzima, apelidada de ***Taq* polimerase**, é muito estável a temperaturas elevadas e não é desnaturada na etapa de preparação da fita na PCR, pode ser adicionada à mistura reacional no início do processo de PCR e continuará funcionando por muitos ciclos.

A segunda inovação fundamental foi o desenvolvimento de termocicladores automáticos, máquinas que produzem as mudanças rápidas de temperatura necessárias para as diferentes etapas de PCR. Originalmente, os tubos com as misturas reacionais são deslocados à mão entre os banhos-maria ajustados em diferentes temperaturas necessárias para as três etapas de cada ciclo. Nos termocicladores automáticos, os tubos de reação são colocados em um bloco de metal que muda a temperatura de acordo com um programa de computador.

Além de amplificar o DNA, a PCR pode ser usada para amplificar as sequências correspondentes para RNA. Essa amplificação é feita ao converter o RNA em DNA complementar (cDNA) com o uso da enzima transcriptase reversa. O cDNA pode ser então amplificado pela PCR comum. Esta técnica é chamada de **PCR de transcrição reversa**.

Limitações da PCR. A reação em cadeia da polimerase é usada atualmente no lugar da clonagem de genes, mas tem muitas limitações. Primeiro, o uso de PCR requer o conhecimento prévio de pelo menos parte da sequência do DNA-alvo para permitir a construção dos *primers*. Segundo, a capacidade da PCR em amplificar quantidades muito pequenas de DNA torna a contaminação um problema grave. Quantidades mínimas de DNA da pele dos funcionários do laboratório ou pequenas partículas no ar podem entrar no tubo de reação e ser amplificadas com o DNA-alvo. São necessários técnica laboratorial cuidadosa e o uso de controles para contornar este problema.

Uma terceira limitação da PCR é a acurácia. Diferente das outras DNA polimerases, a *Taq* polimerase não tem a capacidade de revisar (ver *Capítulo 12*) e sob as condições da PCR, ela incorpora um nucleotídio incorreto a cada 20.000 pb. Foram isoladas novas DNA polimerases estáveis sob calor com a capacidade de revisar, dando mais precisão aos resultados da PCR.

Uma quarta limitação é que o tamanho dos fragmentos que podem ser amplificados pela *Taq* polimerase padrão é em geral de menos de 2.000 pb. Ao usar uma combinação de *Taq* polimerase e uma DNA polimerase com capacidade de revisar, e ao modificar as condições reacionais, os investigadores conseguiram estender a amplificação da PCR para fragmentos maiores, mas até estes fragmentos estão limitados a um comprimento de 50.000 pb ou menos.

Aplicações da PCR. Apesar de suas limitações, a PCR se tornou uma das ferramentas mais usadas na biologia molecular. Os *primers* usados na PCR são específicos para sequências de DNA conhecidas, então a PCR pode ser usada para detectar a uma sequência específica de DNA em uma amostra. Por exemplo, a PCR é usada para detectar o vírus em amostras de sangue ao adicionar *primers* complementares a sequências de DNA virais conhecidas à reação. Se houver DNA viral, os *primers* se ligarão a este e a PCR amplificará um fragmento de DNA de tamanho conhecido. O achado de um fragmento de DNA de tamanho adequado em um gel indica a existência de DNA viral na amostra de sangue. Os modernos testes de diagnóstico para infecção pelo HIV, o agente responsável pela AIDS (ver *Capítulo 9*), usam este tipo de amplificação da PCR das sequências do HIV.

Outra aplicação comum da PCR é identificar a variação genética em populações naturais. A PCR pode ser usada para amplificar segmentos específicos de DNA, que são então analisados por outras ferramentas moleculares que identificam diferenças nos segmentos de nucleotídios. Também, os *primers* de uma variante específica de DNA podem ser usados para determinar se esta variante existe no genoma de um organismo.

A PCR também permite o isolamento do DNA de fontes ancestrais, como DNA de neandertais (discutido na introdução do *Capítulo 10*). É comum usar esta ferramenta para amplificar pequenas quantidades de DNA de cenas de crime e identificar a fonte de uma amostra de DNA por meio de detecção da variação microssatélite (ver seção sobre *fingerprinting* do DNA no final deste capítulo).

A PCR também pode ser usada para introduzir novas sequências em um fragmento de DNA. Embora a extremidade 3′ dos *primers* usados na PCR deva ser complementar com o molde, pode existir uma flexibilidade nas sequências na extremidade 5′ do *primer*. Os *primers* podem ser projetados de modo que tenham novos sítios de restrição ou outras sequências desejáveis nas suas extremidades 5′. Estas sequências então são copiadas pela reação e então adicionadas às extremidades do DNA amplificado por PCR.

Uma modificação da PCR, conhecida como **PCR em tempo real**, pode ser usada para a determinação quantitativa do ácido

nucleico inicial. Nesse procedimento, a PCR normal é usada para amplificar um fragmento específico de DNA e um instrumento sensível é usado para determinar com precisão a quantidade de DNA na solução após cada ciclo de PCR. Uma sonda que fluoresce específica para a sequência de DNA de interesse é usada frequentemente na reação, e então apenas o DNA de interesse é medido. A técnica é chamada de PCR em tempo real porque a quantidade de DNA amplificado é determinada à medida que a reação prossegue.

A PCR em tempo real é combinada com a PCR de transcriptase reversa para quantificar o mRNA em uma amostra, permitindo que os biólogos determinem o nível de expressão gênica em diferentes células e em diferentes condições. Por exemplo, os pesquisadores interessados em como a expressão gênica muda em resposta à administração de um fármaco usam a PCR em tempo real para medir a quantidade de mRNA produzida por genes específicos em células expostas ao fármaco e compará-la com a quantidade de mRNA produzida pelos genes nas células de controle sem exposição ao fármaco.

Conceitos

A reação em cadeia da polimerase é um método enzimático, *in vitro*, para amplificar rapidamente o DNA. Neste processo, o DNA é aquecido para separar as duas fitas, *primers* curtos se fixam ao DNA-alvo e a DNA polimerase sintetiza novas fitas de DNA a partir dos *primers*. Cada ciclo de PCR dobra a quantidade de DNA. A PCR tem várias aplicações importantes na biologia molecular.

Checagem dos conceitos 5

Por que é importante usar uma DNA polimerase termoestável na PCR?

19.3 As técnicas moleculares podem ser usadas para encontrar genes de interesse

Para analisar um gene ou transferi-lo para outro organismo, primeiro é crucial localizá-lo e isolá-lo. Por exemplo, se quisermos transferir um gene humano do hormônio do crescimento para bactérias, primeiro precisamos encontrar o gene que codifica este hormônio e separá-lo dos 3,2 bilhões de pb do DNA humano. Até o momento, em nossa análise de clonagem de genes, contornamos o problema de encontrar a sequência de DNA a ser clonada; a solução para este problema foi postergada de propósito até o momento porque, paradoxalmente, os pesquisadores precisam com frequência clonar um gene para encontrá-lo.

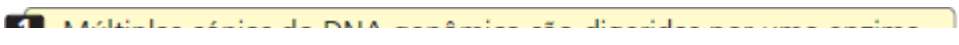
Essa abordagem, clonar primeiro e procurar depois, é chamada de clonagem *shotgun*, porque é como caçar com uma espingarda: as balas se dispersam na direção geral da caça, com boa chance de uma ou mais balas acertarem o alvo em questão. Nessa abordagem, primeiro, um pesquisador clona vários fragmentos de DNA, sabendo que um ou mais têm o DNA de interesse, e então procura o fragmento de interesse entre os clones.

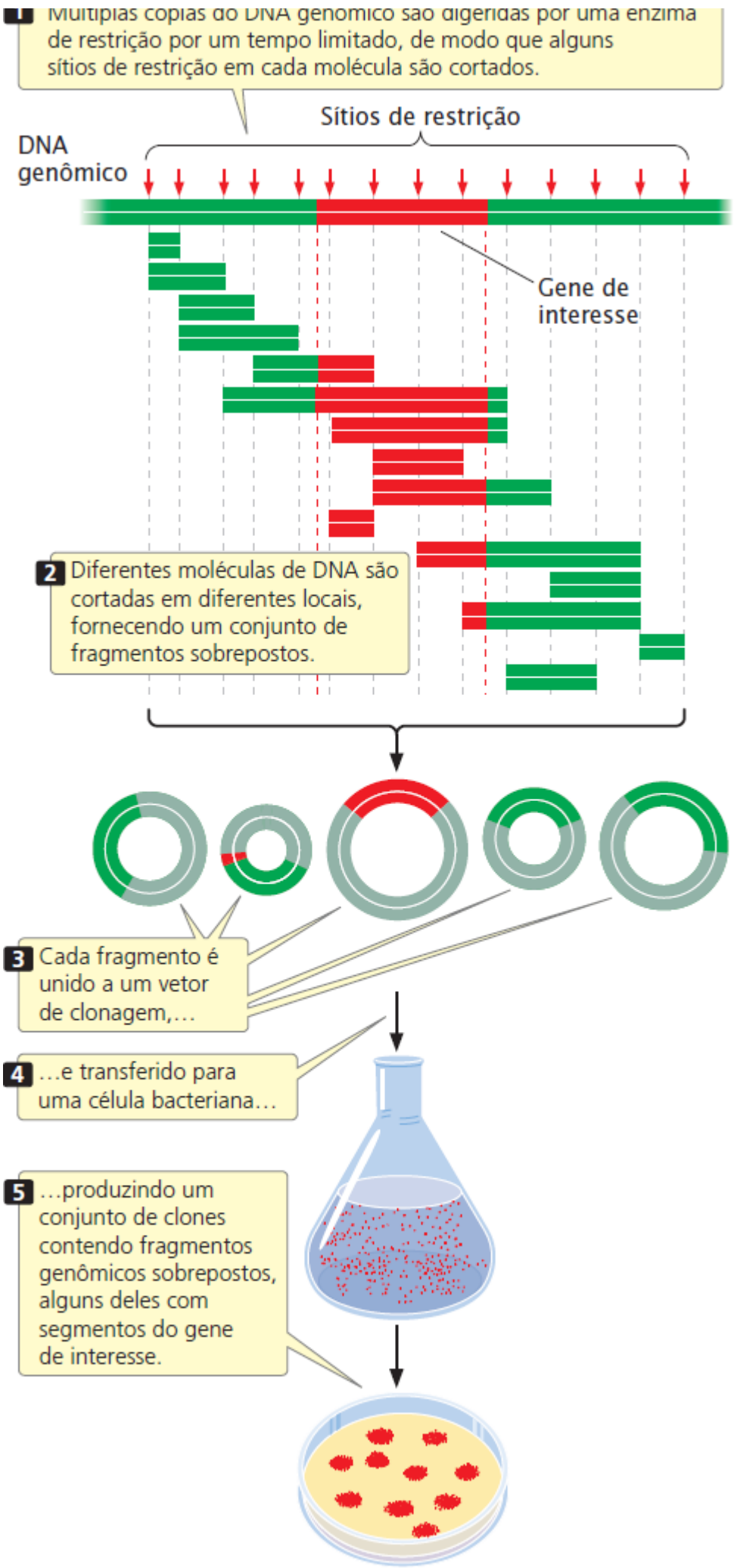
Bibliotecas gênicas

Uma coleção de genes contendo todos os fragmentos de DNA de uma fonte é chamado de **biblioteca de DNA**. Por exemplo, poderíamos isolar o DNA genômico das células humanas, fragmentá-lo, inserir os fragmentos em vetores e cloná-los nas células bacterianas. O conjunto de colônias bacterianas ou fagos contendo estes fragmentos é uma **biblioteca genômica** humana, com todas as sequências de DNA encontradas no genoma humano.

Como criar uma biblioteca genômica. Para criar uma biblioteca genômica, as células são coletadas e rompidas, o que faz com que liberem seu DNA e outras partes do seu conteúdo em uma solução aquosa e o DNA é extraído da solução. Após o DNA ter sido isolado, ele é fracionado por uma enzima de restrição durante um período de tempo limitado (digestão parcial) de modo que apenas *alguns* sítios de restrição em cada molécula de DNA sejam cortados. Como o corte dos sítios é aleatório, diferentes moléculas de DNA serão cortadas em diferentes locais e será produzido um conjunto de fragmentos sobrepostos (**Figura 19.12**). Os fragmentos são então unidos aos vetores, que podem ser transferidos para as bactérias. Alguns clones têm o gene de interesse inteiro (se o gene não for muito grande) e alguns têm partes do gene, mas a maioria tem fragmentos que não são parte do gene de interesse.

Uma biblioteca genômica precisa ter um grande número de clones para garantir que todas as sequências de DNA no genoma estejam representadas na biblioteca. Uma biblioteca do genoma humano formado ao usar cosmídios, cada um carreando um fragmento de DNA aleatório de 35.000 a 44.000 pb de comprimento, precisaria de 350.000 clones de cosmídio para fornecer uma chance de 99% de que cada sequência seja incluída na biblioteca.




1 Múltiplas cópias do DNA genômico são digeridas por uma enzima de restrição por um tempo limitado, de modo que alguns sítios de restrição em cada molécula são cortados.
Sítios de restrição
DNA genômico
Gene de interesse
2 Diferentes moléculas de DNA são cortadas em diferentes locais, fornecendo um conjunto de fragmentos sobrepostos.
3 Cada fragmento é unido a um vetor de clonagem,...
4 ...e transferido para uma célula bacteriana...
5 ...produzindo um conjunto de clones contendo fragmentos genômicos sobrepostos, alguns deles com segmentos do gene de interesse.

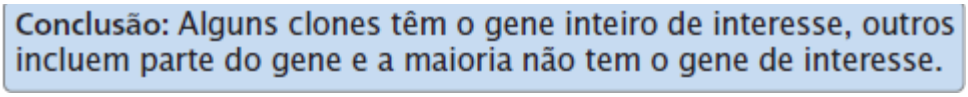


Figura 19.12 Uma biblioteca genômica tem todas as sequências de DNA encontradas em um genoma de um organismo.

Bibliotecas de cDNA. Uma alternativa para criar uma biblioteca genômica é criar uma biblioteca composta apenas por sequências de DNA que são transcritas em mRNA (chamada de **biblioteca de cDNA** porque todo DNA nesta biblioteca é *complementar* ao mRNA). Boa parte do DNA eucariótico é composto por sequências repetitivas (e outras de DNA) que não são transcritas em mRNA (ver *Capítulo 11*) e estas sequências não são representadas em uma biblioteca de cDNA.

Uma biblioteca de cDNA tem duas vantagens adicionais. Primeiro, é enriquecida com fragmentos de genes ativamente transcritos. Segundo, íntrons não interrompem as sequências clonadas; eles são um transtorno quando o objetivo é produzir uma proteína eucariótica nas bactérias, porque a maioria das bactérias não tem meios para remover os íntrons.

A desvantagem da biblioteca de cDNA é que ela tem apenas sequências que estão presentes no mRNA maduro. Às vezes, os pesquisadores estão interessados em sequências que não são transcritas, como as existentes nos promotores e acentuadores, que são importantes para transcrição, mas não são transcritas. Essas sequências não estão presentes em uma biblioteca de cDNA. Além disso, esta biblioteca tem apenas as sequências de genes expressas no tecido a partir do qual o RNA foi isolado e a frequência de determinada sequência de DNA em uma biblioteca de cDNA depende da abundância do mRNA correspondente no tecido em questão. Então, se um gene específico não for expresso ou for expresso apenas em uma frequência baixa em determinado tecido, ele pode estar ausente na biblioteca de cDNA preparada a partir desse tecido. Ao contrário, quase todos os genes são encontrados na mesma frequência em uma biblioteca genômica de DNA.

Para criar uma biblioteca de cDNA, primeiro o RNA mensageiro tem de ser separado de outros tipos de RNA celular (tRNA, rRNA, snRNA etc.). A maioria dos mRNAs eucarióticos tem uma cauda poli(A) que ajuda a separar o mRNA eucariótico dos outros tipos. O RNA celular total é isolado das células e vertido em uma coluna preenchida com fragmentos curtos de DNA compostos inteiramente de nucleotídios timina, ou seja, cadeias oligo(dT) (**Figura 19.13 A**). À medida que o RNA se desloca pela coluna, as caudas poli(A) das moléculas de mRNA pareiam com as cadeias de oligo(dT) e são retidas na coluna, enquanto o resto do RNA passa. O mRNA pode então ser lavado da coluna pela adição de um tampão que rompe as pontes de hidrogênio entre as caudas poli(A) e as cadeias oligo (dT).

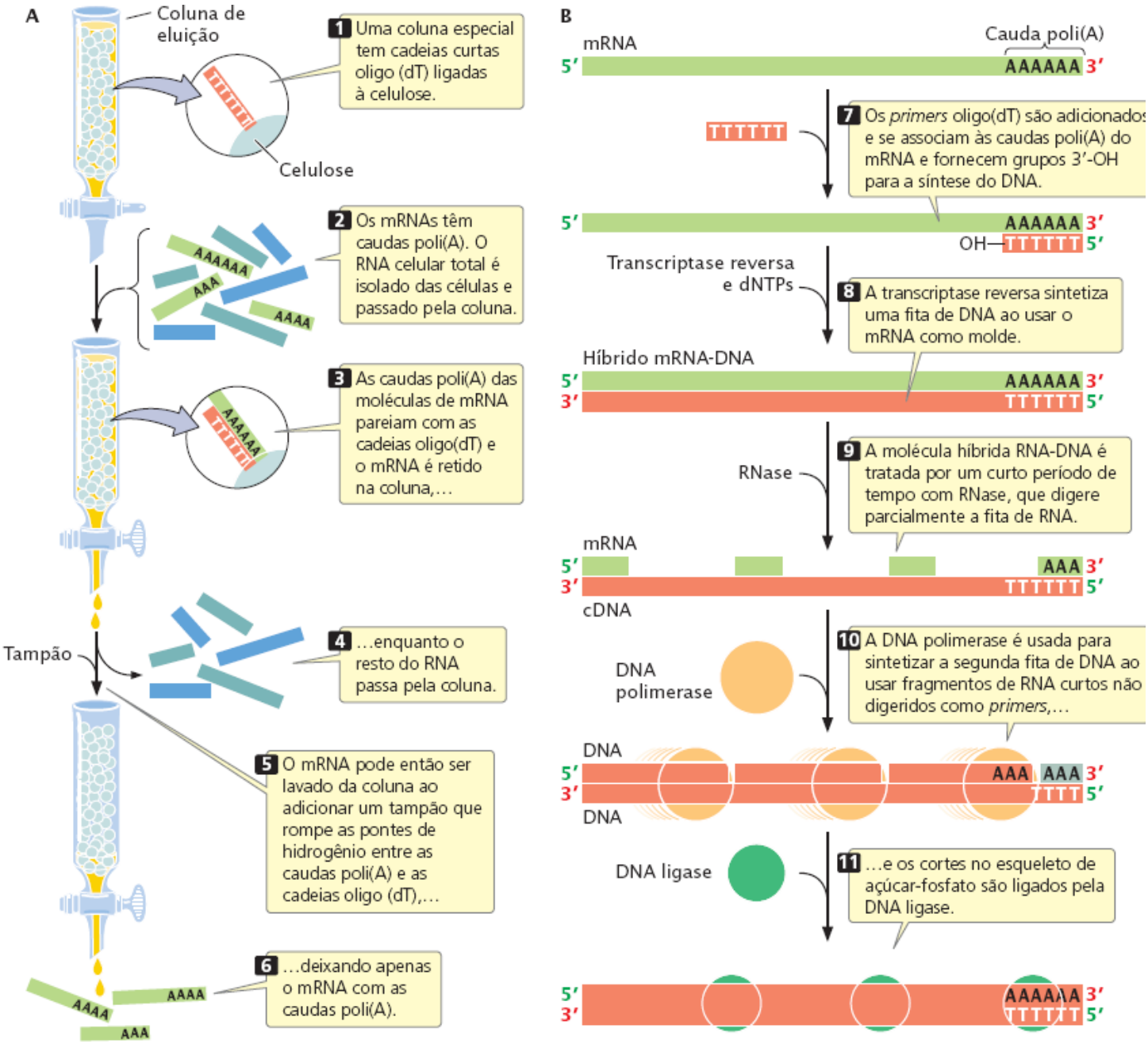


Figura 19.13 Uma biblioteca de cDNA tem apenas as sequências de DNA que são transcritas em mRNA.

As moléculas de mRNA são então copiadas em cDNA. A transcriptase reversa, uma enzima isolada dos retrovírus (ver *Capítulo 9*), sintetiza o DNA complementar de fita única a partir do molde de RNA (**Figura 19.13 B**). O híbrido RNA-DNA resultante é finalmente convertido em uma molécula de cDNA de fita dupla pela DNA polimerase. Resolva o Problema 35

Conceitos

Um método para encontrar um gene é criar e rastrear uma biblioteca de DNA. Uma biblioteca genômica é criada pelo corte do DNA genômico em fragmentos sobrepostos e pela clonagem de cada fragmento em uma célula bacteriana separada. Uma biblioteca de cDNA é criada a partir do mRNA que é convertido em cDNA e clonado nas bactérias.

Como rastrear bibliotecas de DNA. Criar uma biblioteca genômica ou de cDNA é relativamente fácil se comparado com a seleção da biblioteca para encontrar clones com o gene de interesse. O procedimento de seleção usado depende do que se sabe sobre o gene.

A primeira etapa no rastreamento é semear (plaquear) os clones da biblioteca. Se for usado um plasmídio ou cosmídio para construir a biblioteca, as células são diluídas e semeadas de forma que cada bactéria cresça em uma colônia distinta. Se for usado um fago como vetor, os fagos são usados para infectar uma camada de bactérias em uma placa de Petri. Cada placa ou colônia bacteriana tem um único fragmento de DNA clonado que tem de ser rastreado à procura do gene de interesse.

Um modo comum de rastreamento das bibliotecas é usar sondas. Para usar uma sonda, primeiro é necessário fazer réplicas das colônias ou das placas semeadas na biblioteca. A **Figura 19.14** ilustra esse procedimento para uma biblioteca de cosmídio.

Como uma sonda é obtida quando um gene ainda não foi isolado? Uma opção é usar um gene semelhante oriundo de outro organismo como sonda. Por exemplo, se quisermos rastrear uma biblioteca genômica humana à procura do gene do hormônio do crescimento e o gene já foi isolado de ratos, podemos usar uma sequência gênica purificada de rato como sonda para encontrar o gene humano para esse hormônio. Uma hibridização bem-sucedida não demanda complementaridade perfeita entre a sonda e a sequência-alvo, então uma sequência semelhante pode ser usada como uma sonda.

Por outro lado, podem ser criadas sondas sintéticas se a proteína produzida pelo gene já foi isolada e sua sequência de aminoácidos já foi determinada. Com o uso do código genético e da sequência de aminoácidos da proteína, podem ser deduzidas possíveis sequências de nucleotídios de uma pequena região do gene. Embora apenas uma sequência no gene codifique uma proteína específica, a existência de códons sinônimos significa que a mesma proteína poderia ser produzida por diferentes sequências de nucleotídios e é impossível saber qual está correta. Para superar este problema, é usada uma mistura de todas as possíveis sequências de nucleotídios como sonda. Para minimizar o número de sequências necessárias na mistura, é selecionada uma região da proteína com relativamente pouca degeneração nos seus códons. Quando parte da sequência de DNA do gene for determinada, um conjunto de sondas de DNA pode ser sintetizado quimicamente por uma máquina automática chamada de sintetizador de oligonucleotídios.

Outro método de rastreamento de uma biblioteca é procurar pelo produto de um gene, a proteína. A biblioteca de cDNA precisa ser clonada em um vetor de expressão neste método. Os clones podem ser testados à procura da proteína ao usar um anticorpo que reconheça a proteína ou ao usar um teste químico para a proteína. Este método depende da existência de um teste para a proteína produzida pelo gene. As bibliotecas gênicas também podem ser rastreadas usando PCR ou sequenciamento.

Conceitos

Uma biblioteca de DNA pode ser rastreada à procura de um gene específico por sondas complementares que hibridizam no gene. Por outro lado, a biblioteca pode ser clonada em um vetor de expressão e o gene pode ser localizado ao se examinarem os clones à procura da proteína do gene.

Checagem dos conceitos 6

Explique resumidamente como as sondas sintéticas são criadas para rastrear uma biblioteca de DNA quando a proteína codificada pelo gene é conhecida.

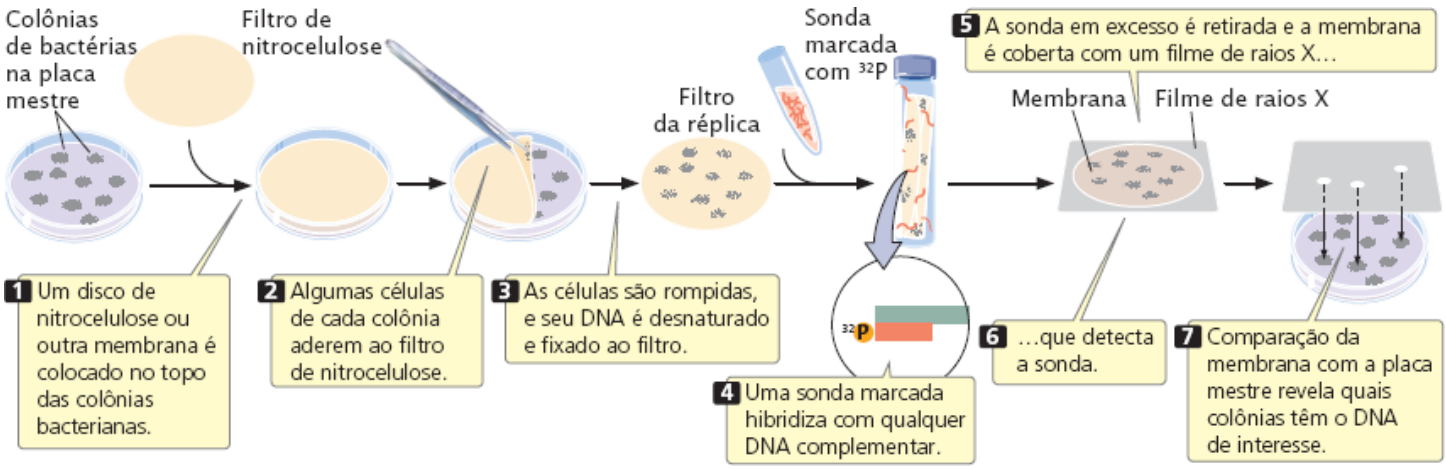


Figura 19.14 **As bibliotecas genômicas e de cDNA podem ser rastreadas com uma sonda para encontrar o gene de interesse.**

Hibridização *in situ*

As sondas de DNA podem ser usadas para determinar a localização cromossômica de um gene em um processo chamado de **hibridização *in situ***. O nome é derivado do fato de que o DNA (ou RNA) é visualizado enquanto está na célula (*in situ*). Esta técnica requer que as células sejam fixadas e os cromossomos sejam espalhados em uma lâmina de microscópio e desnaturados. Uma sonda marcada é então aplicada à lâmina, como é aplicada ao gel. Muitas sondas carreiam corantes fluorescentes que podem ser observados diretamente ao microscópio (**Figura 19.15 A**). Várias sondas com diferentes corantes coloridos podem ser usadas simultaneamente para investigar diferentes sequências ou cromossomos. A hibridização *in situ* com fluorescência (FISH) é amplamente usada para identificar a localização cromossômica dos genes humanos.

Esse método também pode ser usado para determinar a distribuição de moléculas de mRNA específicas no tecido, servindo como fonte de sugestão sobre as distinções da expressão gênica entre diferentes células (**Figura 19.15 B**). Uma sonda de DNA marcada complementar a uma molécula de mRNA específica é adicionada ao tecido, e a localização da sonda é determinada com o uso de marcadores radioativos ou fluorescentes. Determinar onde um gene é expresso ajuda a definir sua função. Por exemplo, descobrir que um gene é muito expresso apenas no tecido do cérebro pode sugerir que este gene participe na função neurológica.

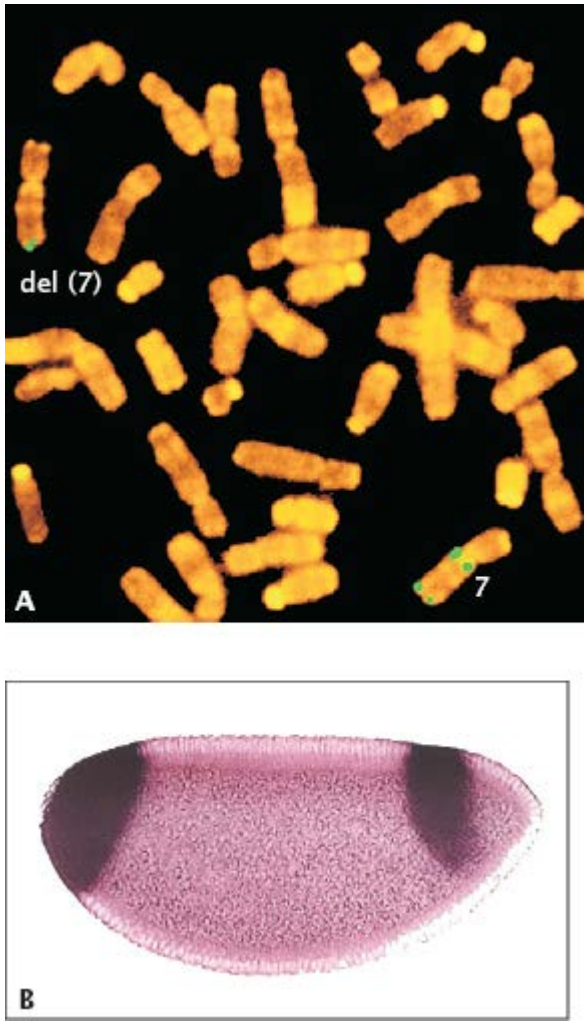


Figura 19.15 Com a hibridização *in situ*, as sondas de DNA são usadas para determinar a localização celular ou cromossômica de um gene ou seu produto. A. Uma sonda com fluorescência verde é específica para o cromossomo 7, revelando uma deleção em uma cópia do cromossomo 7. **B.** A hibridização *in situ* é usada para detectar mRNA oriundo do gene *tailless* em um embrião de *Drosophila*. (*Parte A*: Addenbrookes Hospital/Science Source. *Parte B*: Cortesia de L. Tsuda.)

Clonagem posicional

Para muitos genes com funções importantes, nenhuma proteína associada é conhecida até o presente. As bases bioquímicas de muitas doenças genéticas humanas, por exemplo, ainda são desconhecidas. Como estes genes podem ser isolados? Uma forma é determinar primeiro a localização geral do gene no cromossomo ao usar as frequências de recombinação derivadas de cruzamentos ou linhagens (ver *Capítulo 7*). Após ser identificada a região cromossômica onde o gene está, os genes nessa região podem ser clonados e identificados. Outras técnicas podem ser usadas para identificar qual dos genes "candidatos" pode ser o que provoca a doença. Esta abordagem, isolar genes com base na sua posição em um mapa gênico, é chamada de **clonagem posicional**.

Na primeira etapa da clonagem posicional, os geneticistas usam estudos de mapeamento (ver *Capítulo 7*) para estabelecer a ligação entre os marcadores moleculares e um fenótipo de interesse, como a doença humana ou um traço físico desejável em uma planta ou animal. A demonstração de ligação entre o genótipo e um ou mais marcadores moleculares forneceria informações sobre qual cromossomo carreia o *locus* que codifica o fenótipo e sua localização geral naquele cromossomo.

A próxima etapa é localizar com maior precisão o *locus* ao usar marcadores moleculares adicionais agrupados na região cromossômica onde está o *locus*. Após o gene ser colocado em um mapa cromossômico, os clones que cobrem a região podem ser isolados a partir de uma biblioteca genômica. Com o uso de uma técnica chamada **caminhada cromossômica** (**Figura**

19.16), é possível progredir a partir dos genes vizinhos para os clones relacionados, um dos quais pode ter o gene de interesse. A base da caminhada cromossômica é o fato de que uma biblioteca genômica é composta por um conjunto de fragmentos de DNA *sobrepostos* (ver Figura 19.12). Começamos com um marcador de gene clonado que está próximo do novo gene de interesse, então a "caminhada" será tão curta quanto possível. Uma extremidade do clone de um marcador vizinho (clone A na Figura 19.16) é usada para fazer uma sonda complementar. Essa sonda é usada para rastrear a biblioteca genômica à procura de um segundo clone (clone B) que se sobrepõe ao primeiro e se estende na direção do gene de interesse. Este segundo clone é isolado e purificado e uma sonda é preparada a partir da sua extremidade. A segunda sonda é usada para rastrear a biblioteca para um terceiro clone (clone C) que se sobrepõe com o segundo. Desta forma, pode-se fazer a "caminhada" sistematicamente na direção do gene de interesse, um clone por vez.

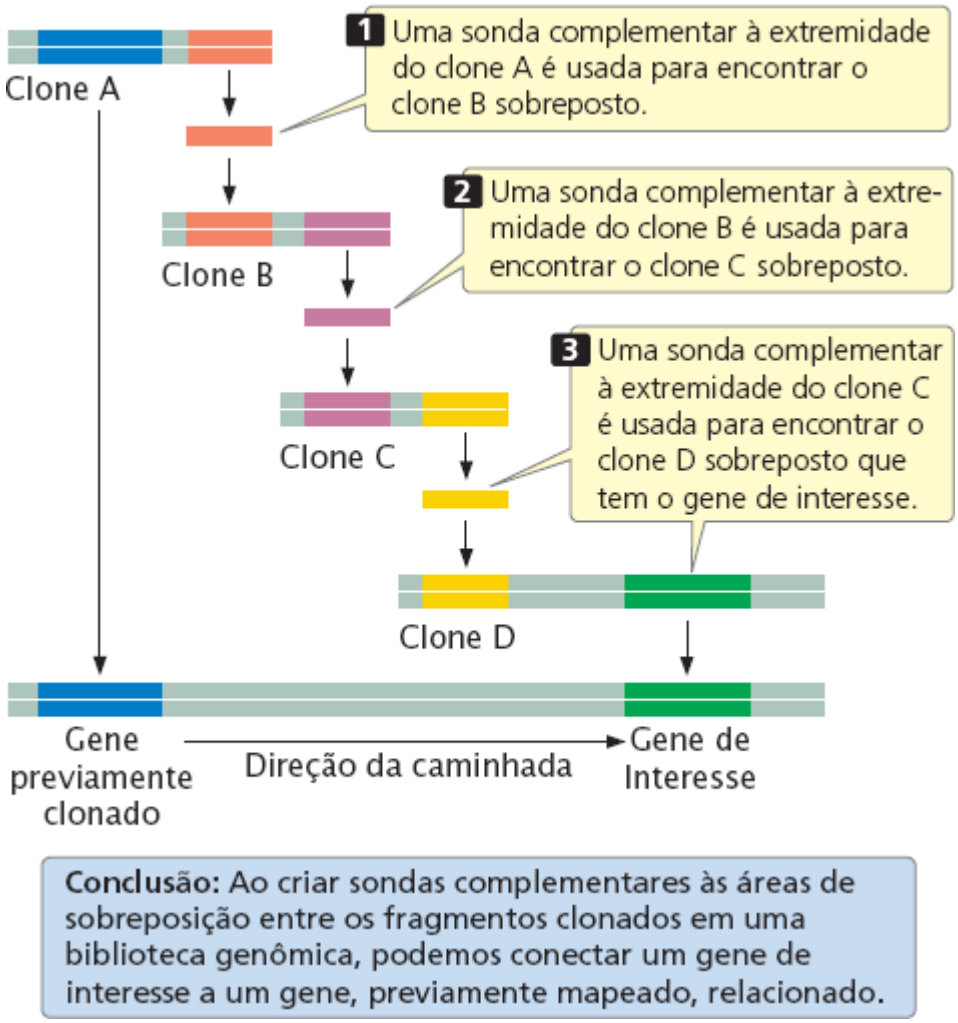


Figura 19.16 Na caminhada cromossômica, os genes vizinhos são usados para localizar um gene de interesse.

Uma técnica relacionada chamada salto cromossômico permite que se mova a partir de marcadores ligados mais distantes a clones que têm uma sequência de interesse. Após os clones que cobrem a região delineada serem obtidos por caminhada ou salto cromossômico, todos os genes localizados dentro da região são identificados. Os genes podem ser diferenciados de outras sequências pela presença de recursos característicos, como as sequências consenso no promotor e um códon de início e um códon de parada dentro do mesmo quadro de leitura. Após os genes "candidatos" serem identificados, eles podem ser avaliados para determinar qual tem a maior probabilidade de ser o gene de interesse. O padrão de expressão do gene, onde e quando ele é transcrito, pode oferecer dicas sobre seu funcionamento. Por exemplo, é provável que os genes para doença neurológica sejam expressos no cérebro. Os geneticistas procuram mutações na região codificadora do gene entre as pessoas com a doença. As próximas seções e o Capítulo 20 trazem mais informações sobre a determinação da função dos genes.

Conceitos

A clonagem posicional permite que os pesquisadores isolem um gene sem conhecer sua base bioquímica. Os estudos de ligação são usados para mapear o *locus* que produz um fenótipo de interesse para uma região cromossômica específica. A caminhada e o salto cromossômicos podem ser usados para progredir a partir dos marcadores moleculares a clones contendo as sequências que cobrem a região do cromossomo. Os

genes candidatos dentro da região são avaliados para determinar se eles codificam o fenótipo de interesse.

Checagem dos conceitos 7

Como os genes candidatos identificados pela clonagem posicional são avaliados para saber se eles codificam o fenótipo em questão?

Aplicação | Isolamento do gene para a fibrose cística

O primeiro gene responsável por uma doença genética humana que foi isolado inteiramente pela clonagem posicional foi o gene responsável pela fibrose cística (FC). A fibrose cística é um distúrbio autossômico recessivo caracterizado por infecções crônicas dos pulmões, produção insuficiente de enzimas pancreáticas necessárias para digestão e aumento de concentração de sal no suor. Está entre as doenças genéticas mais comuns nos caucasianos, com uma frequência de 1 em cada 2.000 nascidos vivos. Quase 5% de todos os caucasianos são carreadores da mutação da FC.

Os geneticistas que tentaram isolar o gene para FC lidaram com uma tarefa formidável. Os sintomas da doença, em especial a elevada concentração de sal no suor, sugerem que o gene para FC de alguma forma participe dos movimentos dos íons para dentro e para fora da célula, mas não existia informação disponível sobre a proteína codificada pelo gene. Naquele momento o genoma humano ainda não havia sido sequenciado. As análises de linhagem mostraram que a FC é herdada como um traço autossômico recessivo e então ela pode estar localizada em qualquer um dos 22 pares de cromossomos autossômicos. Assim, os geneticistas estavam procurando um gene desconhecido, provavelmente incluindo milhares ou dezenas de milhares de pares de bases, entre os 3,2 bilhões de pares de bases do genoma humano.

Os pesquisadores começaram a procurar por associações entre a hereditariedade da FC e de outros traços (**Figura 19.17**). Os estudos iniciais eram limitados pela escassez de traços genéticos que variaram e poderiam ser usados para estudos de mapeamento gênico, mas nos anos de 1980, os avanços na biologia molecular forneceram vários marcadores moleculares que poderiam ser usados para a análise de ligação (ver *Capítulo 7*). Os geneticistas coletaram as linhagens de famílias que tinham vários membros com FC. Eles comparavam a hereditariedade da FC com de outros marcadores moleculares entre os membros destas famílias, procurando por indício de ligação. Foi descoberto que o gene para FC estava intimamente relacionado com dois marcadores, MET e D7S8, localizados no braço longo do cromossomo 7. MET e D7S8 estão separados em cerca de 1,5 unidade de mapa (ver *Capítulo 7*). No genoma humano, cada unidade de mapa corresponde a aproximadamente 1 milhão de pares de bases, então o gene para FC está localizado em algum lugar dentro de um segmento de 1,5 milhão de pares de bases de DNA, uma ampla expansão de sequência.

Outros estudos de ligação com marcadores adicionais foram feitos para delinear com maior precisão onde o gene para FC reside na região de 1,5 milhão de pares de base. Os pesquisadores selecionaram marcadores moleculares adicionais a partir da região circundando MET e D7S8 e fizeram estudos de ligação entre estes novos marcadores e a FC (ver Figura 19.17). Estes estudos identificaram dois marcadores adicionais, D7S122 e D7S340, que estão intimamente ligados a FC. Além disso, eles mostraram que a ordem dos quatro marcadores é MET-D7S340-D7S122-D7S8 e que o gene FC está muito próximo a D7S122 e D7S340. Este achado estreitou a região na qual o gene para FC está em cerca de 500.000 pb.

Nesse estágio, os geneticistas começaram a isolar os clones das sequências a partir da região delimitada. Ao iniciar a partir dos marcadores moleculares, eles usaram uma combinação de caminhada e salto cromossômico para identificar os clones a partir das bibliotecas genômicas humanas que cobrem completamente a região de interesse (ver Figura 19.17). Um exame das sequências dentro destes clones revelou a presença de quatro genes na região comprometida pelos marcadores ligados (ver Figura 19.17). Estudos adicionais foram feitos para caracterizar melhor estes genes candidatos. Três dos genes candidatos foram consequentemente eliminados, porque os estudos de ligação sugeriram que eles não estavam intimamente ligados com a hereditariedade da FC ou por causa da análise das sequências ou seus padrões de expressão sugeriram que eles não eram o gene para FC.

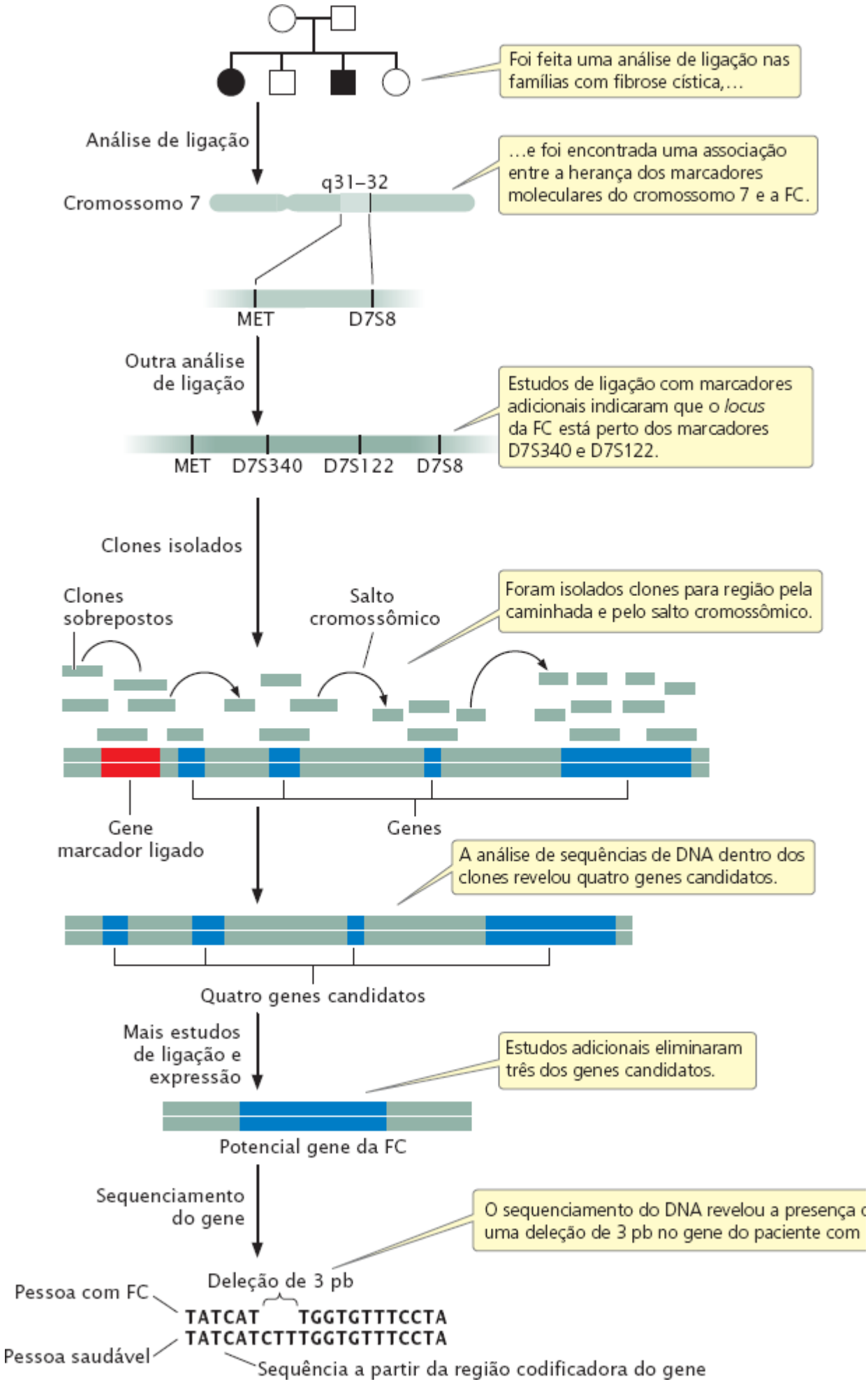


Figura 19.17 **O gene para fibrose cística foi localizado usando a clonagem posicional.**

Foram feitos estudos de hibridização com um gene remanescente para determinar onde ele era expresso. O RNA mensageiro foi isolado a partir de diferentes tecidos de órgãos e sondado com sequências a partir do gene candidato. O gene mostrou altos níveis de expressão no pâncreas, pulmões e glândulas sudoríparas (**Figura 19.18**), tecidos conhecidos por serem afetados pela FC.

Foram sequenciadas as cópias do gene candidato de uma pessoa saudável e uma pessoa com FC e os dados da sequência foram examinados para diferenças que podem ser uma mutação que provoca FC. Os achados revelaram que a pessoa com FC tinha uma deleção de 3 pb na região codificadora do gene, enquanto uma pessoa saudável não tinha esta deleção. A deleção resultava na ausência de um aminoácido fenilalanina da proteína codificada pelo gene candidato. Então, para vários pacientes com FC, os geneticistas usaram PRC para amplificar a região do gene onde foi encontrada a deleção e 68% dos pacientes com FC tinham esta deleção. Estudos posteriores demonstraram que os pacientes com FC remanescentes tinham mutações em outras localizações no gene candidato, provando que o gene candidato era de fato o *locus* responsável pela FC.

Consequentemente, os pesquisadores demonstraram que o gene para a FC codifica uma proteína de membrana que controla o movimento do íon cloro para dentro e para fora das células e é conhecida como o gene *regulador da condutância transmembrana da fibrose cística* (*CFTR*). Pacientes com FC têm duas formas mutantes de *CFTR*, o que faz com que canais de cloreto fiquem fechados. Os íons cloreto presos na célula levam à formação do muco espesso e aos sintomas da doença.

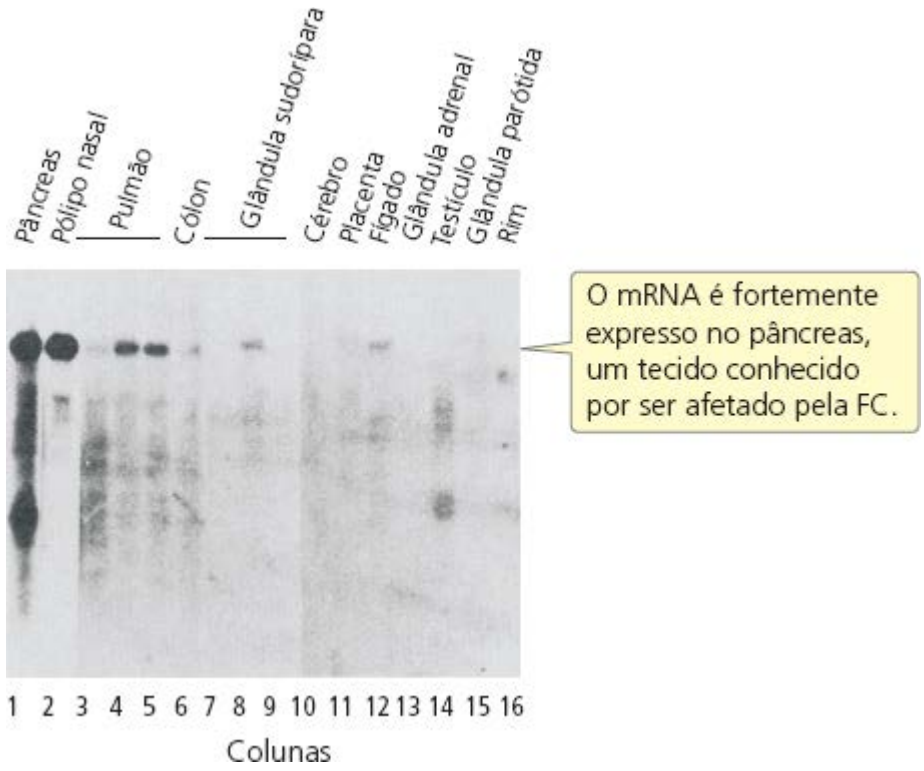


Figura 19.18 Um candidato para o gene para fibrose cística é expresso nos tecidos pancreático, respiratório e das glândulas sudoríparas, tecidos que são afetados pela doença. Esta imagem é um *Northern blot* de mRNA produzido pelo gene candidato nos diferentes tecidos. Estes dados fornecidos evidenciam que o gene candidato é de fato o gene responsável pela fibrose cística. (De J. R. Riordan et al., Science 245:1066-1073, 1989. Reimpressa com permissão de AAAS.)

19.4 As sequências podem ser determinadas e analisadas

Além de clonar e amplificar o DNA, as técnicas moleculares são usadas para analisar as moléculas de DNA pelo estudo e pela determinação de suas sequências.

Polimorfismos de comprimento do fragmento de restrição

Uma contribuição significativa da genética molecular é fornecer numerosos marcadores genéticos que podem ser usados no mapeamento gênico. Aprendemos como estes marcadores são essenciais para o sucesso da clonagem posicional. Um grupo de tais marcadores inclui os **polimorfismos de comprimento do fragmento de restrição** (RFLPs), que são variações (polimorfismos) nos padrões dos fragmentos produzidos quando as moléculas são cortadas com a mesma enzima de restrição (**Figura 19.19**). Estas diferenças são herdadas e podem ser usadas no mapeamento, semelhante à forma na qual as diferenças alélicas são usadas para mapear genes convencionais.

Para ilustrar o mapeamento com RFLPs, considere a doença de Huntington, um distúrbio autossômico dominante. Na família

apresentada na **Figura 19.20**, o pai é heterozigoto para a doença de Huntington (*Hh*) e para um padrão de restrição (*AC*). A partir do pai, cada criança herda um alelo da doença de Huntington (*H*) ou um alelo normal (*h*); qualquer criança que herda o alelo *H* desenvolve a doença, porque é um distúrbio autossômico dominante. A criança também herda um dos dois alelos RFLP do pai, *A* ou *C*, que produz o padrão RFLP correspondente. Na Figura 19.20, cada criança que herda o padrão *C* do pai também herda a doença de Huntington (e, portanto, o alelo *H*), porque o *locus* para RFLP está intimamente ligado com o *locus* para o gene causador da doença. Se não observarmos correspondência entre a hereditariedade do padrão RFLP e da doença, a falta de correspondência indicaria que os genes que codificam RFLP e a doença de Huntington são classificados separadamente e não estão ligados. Recentemente, a disponibilidade da tecnologia de sequenciamento de DNA com menor custo (ver adiante) reduziu o uso de RFLPs no mapeamento gênico e diagnóstico genético. Atualmente, o diagnóstico da doença de Huntington é feito na rotina usando a ampliação de uma porção do gene da doença com PCR. **Resolva o Problema 41**

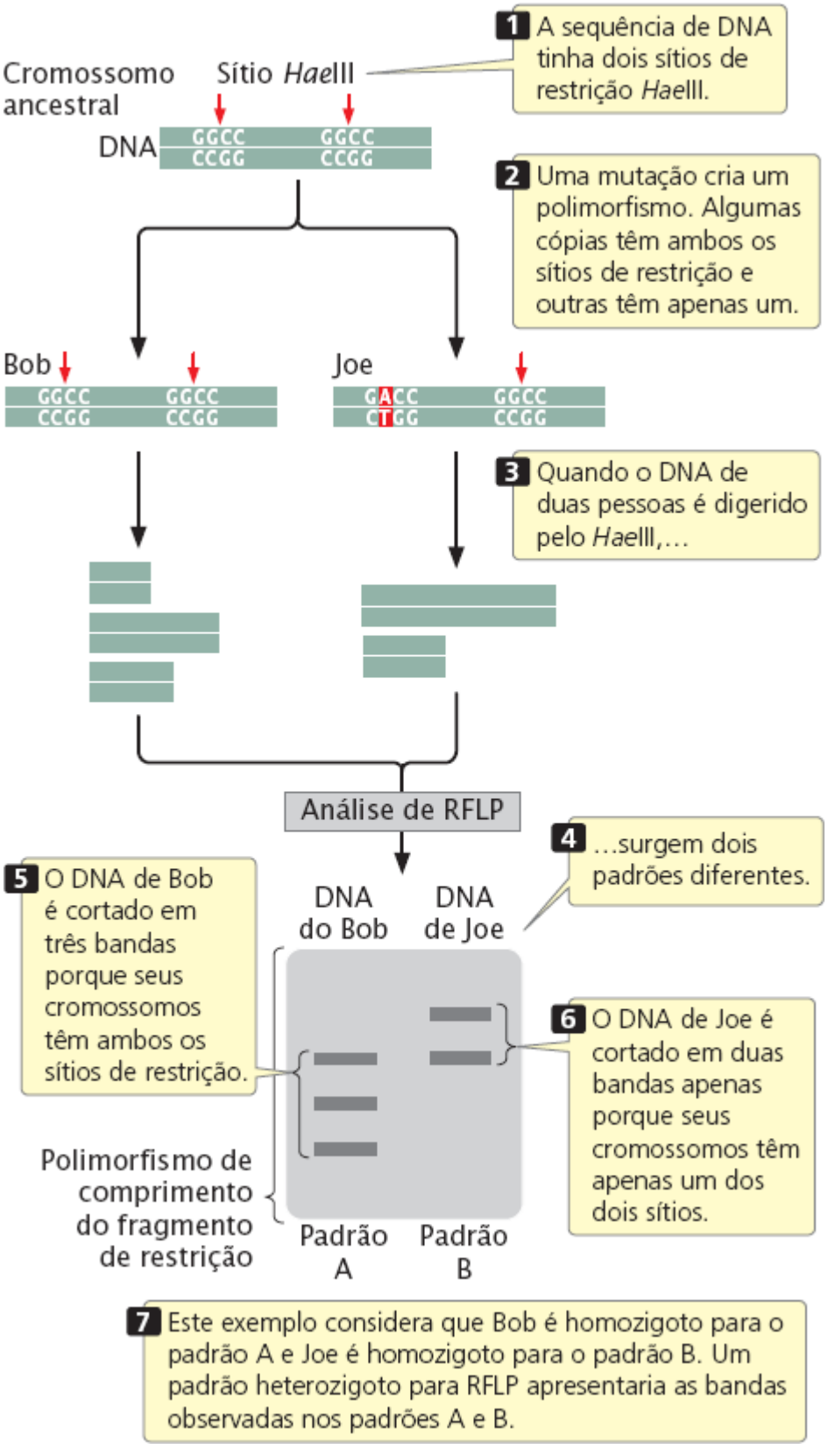


Figura 19.19 Os polimorfismos de comprimento do fragmento de restrição são marcadores genéticos que podem ser usados no mapeamento.

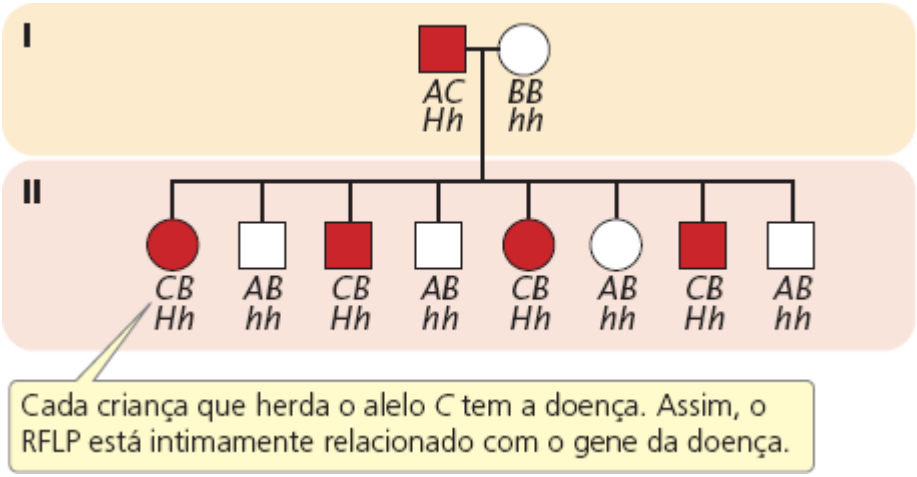


Figura 19.20 Os polimorfismos de comprimento do fragmento de restrição podem ser usados para detectar ligação. A letra *H* refere-se ao alelo da doença de Huntington, o alelo h refere-se ao alelo normal. *A*, *B* e *C* referem-se aos alelos RFLP. Cada criança que herda o alelo *C* do RFLP tem a doença de Huntington, indicando que os genes que codificam o RFLP e a doença de Huntington estão intimamente ligados.

Conceitos

Os polimorfismos de comprimento do fragmento de restrição são variações no padrão dos fragmentos produzidos pelas enzimas de restrição, que revelam variações nas sequências de DNA. Estes polimorfismos são muito usados no mapeamento gênico.

Sequenciamento de DNA

Um método molecular poderoso para analisar DNA é uma técnica conhecida como **sequenciamento de DNA**, que determina de forma rápida a sequência de bases no DNA. O sequenciamento permite que a informação genética no DNA seja lida, fornecendo muita informação sobre a estrutura e função dos genes. Em meados dos anos 1970, Frederick Sanger e seus colegas criaram um método de sequenciamento didesoxi baseado no prolongamento do DNA pela DNA polimerase e, no mesmo período, Allan Maxam e Walter Gilbert desenvolveram um segundo método baseado na degradação química do DNA. O método de Sanger se tornou o procedimento padrão para sequenciamento de qualquer fragmento de DNA purificado.

O método Sanger ou didesoxi de sequenciamento de DNA baseia-se na replicação. O fragmento a ser sequenciado é usado como molde para criar uma série de novas moléculas de DNA. No processo, a replicação é as vezes (mas não sempre) encerrada quando uma base específica é encontrada, produzindo fitas de DNA de diferentes comprimentos, cada qual termina na mesma base.

O método depende do uso de um substrato especial para síntese de DNA. Normalmente, o DNA é sintetizado a partir de trifosfatos de desoxirribonuclesídio (dNTPs), que têm um grupo OH no átomo do carbono 3′ (**Figura 19.21 A**). No método de Sanger, um nucleotídio especial, chamado de **trifosfato de didesoxirribonucleosídio** (ddNTP, **Figura 19.21 B**) é usado como um dos substratos. Os ddNTPs são idênticos aos dNTPs, exceto que eles não têm um grupo 3′-OH. No curso da síntese do DNA, os ddNTPs são incorporados à fita de DNA crescente. Entretanto, após um ddNTP ser incorporado à fita de DNA, não podem ser mais adicionados nucleotídios, porque não existe um grupo 3′-OH para formar uma ligação fosfodiéster com um nucleotídio seguinte. Assim, os ddNTPs encerram a síntese de DNA.

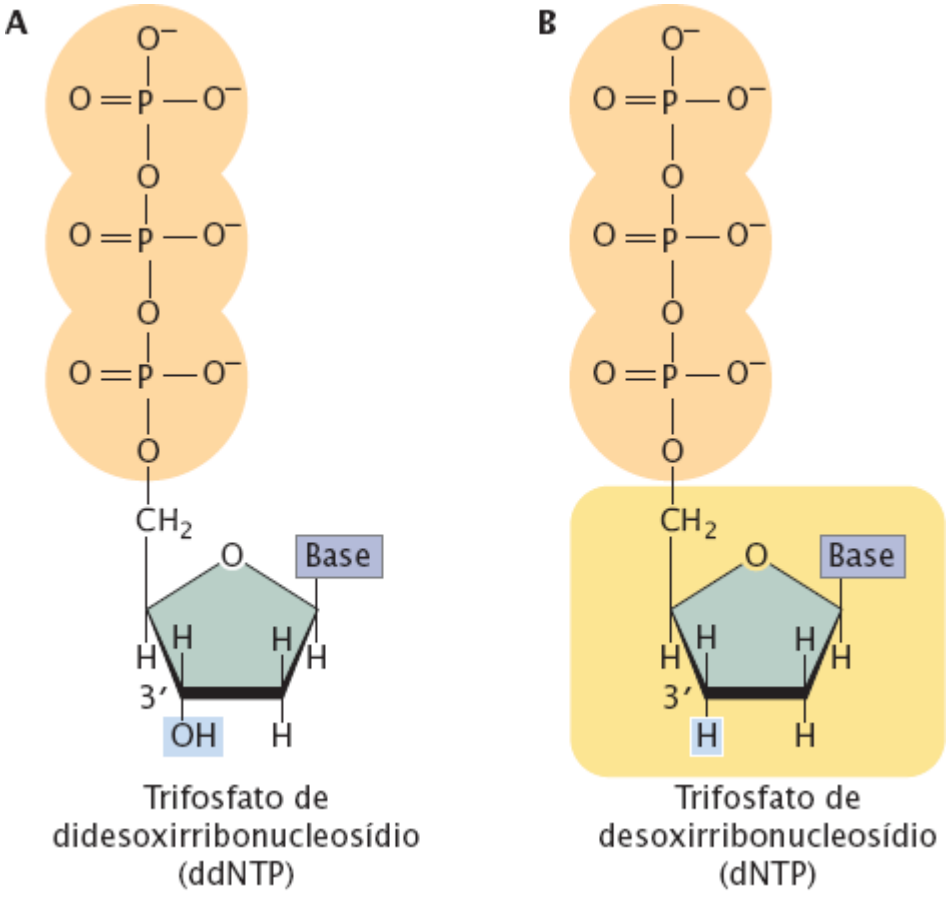


Figura 19.21 A reação de sequenciamento didesoxi requer um substrato especial para a síntese de DNA. A. Estrutura do trifosfato de desoxirribonucleosídio, o substrato normal para síntese de DNA. **B.** Estrutura do trifosfato de didesoxirribonucleosídio, que não tem um grupo OH no átomo do carbono 3′.

Embora seja tecnicamente possível sequenciar uma única molécula de DNA, a maioria dos processos de sequenciamento atualmente requer uma quantidade considerável de DNA, qualquer fragmento de DNA a ser sequenciado deve ser primeiro ampliado por PCR ou por clonagem em bactéria. As cópias do DNA-alvo são isoladas e divididas em quatro partes (**Figura 19.22**). Cada parte é colocada em um diferente tubo, ao qual são adicionados:

1. Várias cópias de um *primer* que é complementar a uma da extremidade da fita do DNA-alvo.
2. Todos os quatro tipos de trifosfatos de desoxirribonucleosídio, os precursores normais da síntese de DNA.
3. Uma pequena quantidade de um dos quatro tipos de trifosfatos de didesoxirribonuclesídio, que terminam a síntese de DNA assim que são incorporados a qualquer cadeia crescente (cada um dos quatro tubos recebendo um ddNTP diferente).
4. DNA polimerase.

O *primer* ou um dos dNTPs é marcado com radioatividade ou substância química de modo que o DNA recém-produzido possa ser detectado.

Dentro de cada um dos quatro tubos, a enzima DNA polimerase sintetiza DNA. Vamos analisar a reação em um dos quatro tubos, um que tenha recebido ddATP. Dentro deste tubo, cada uma das fitas únicas do DNA-alvo serve como molde para síntese de DNA. O *primer* pareia com sua sequência complementar em uma extremidade de cada fita molde, fornecendo um grupo 3′-OH para o início da síntese do DNA. A DNA polimerase prolonga uma nova fita de DNA a partir do seu *primer*. Sempre que a DNA polimerase encontrar um T na fita molde, ela usa aleatoriamente dATP ou ddATP para introduzir um A na fita recém-sintetizada. Como existe mais dATP do que ddATP na mistura reacional, dATP é incorporado com maior frequência, permitindo que a síntese do DNA continue. Em alguns casos, entretanto, o ddATP é incorporado à fita e a síntese termina. A incorporação de ddA à nova fita ocorre de forma aleatória em diferentes posições em diferentes cópias, produzindo um conjunto de cadeias de DNA de diferentes comprimentos (12, 7 e 2 nucleotídios no exemplo ilustrado na Figura 19.22), cada uma terminando em um nucleotídio que tem adenina.

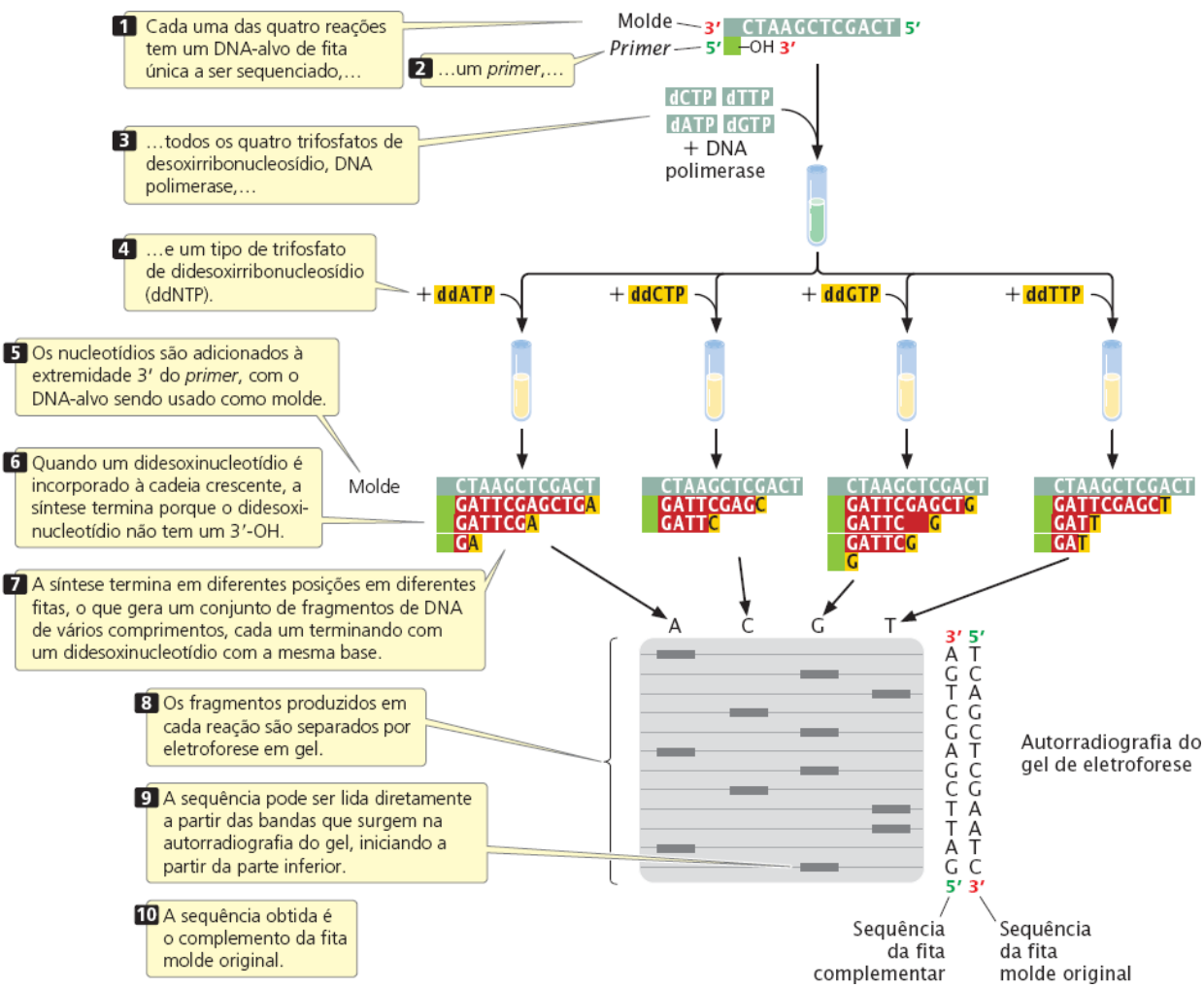


Figura 19.22 O método didesoxi de sequenciamento de DNA baseia-se no término da síntese de DNA.

Reações equivalentes ocorrem nos outros três tubos, exceto que a síntese é terminada em nucleotídios com uma base diferente em cada tubo. Após o término das reações de polimerização, todo DNA nos tubos é desnaturado e os produtos de fita única de cada reação são separados por eletroforese em gel.

O conteúdo dos quatro tubos é separado lado a lado em um gel de acrilamida de modo que as fitas de DNA de diferentes comprimentos com um único nucleotídio possam ser distinguidas. Após a eletroforese, a localização e os tamanhos das fitas de DNA no gel são revelados com o uso de marcadores radioativos ou químicos.

A leitura da sequência de DNA é a parte mais simples e curta do procedimento. Na Figura 19.22, você pode observar a banda mais perto do fundo do gel é do tubo que tinha a reação ddGTP, o que significa que o primeiro nucleotídio sintetizado tinha guanina (G). A próxima banda para cima está no tubo com ddATP, então o próximo nucleotídio na sequência é adenina (A) e assim em diante. Desta forma, a sequência é lida a partir da parte inferior para a parte superior do gel, com os nucleotídios próximos do fundo correspondendo à extremidade 5′ da fita de DNA recém-sintetizada e os que estão próximos do topo correspondem à extremidade 3′. Lembre-se de que a sequência obtida não é a do DNA-alvo, mas do seu *complemento*.

Resolva o Problema 39

Por muitos anos, o sequenciamento do DNA era feito à mão, algo cansativo e caro. Atualmente, o sequenciamento é feito com máquinas automáticas que usam corantes fluorescentes e escâneres a *laser* para sequenciar milhares de pares de base dentro de algumas horas (**Figura 19.23**). A reação didesoxi também é usada aqui, mas os ddNTPs empregados na reação estão marcados com corantes fluorescentes e um diferente corante é usado para cada tipo de didesoxinucleotídio. Neste caso, as quatro reações de sequenciamento podem ocorrer no mesmo tubo de teste e podem ser colocadas no mesmo poço durante a eletroforese. As máquinas para sequenciamento mais recentes fazem a eletroforese em tubos capilares com gel. Os fragmentos de diferentes tamanhos produzidos pela reação de sequenciamento separam-se dentro do tubo e migram para um feixe de *laser* e detector. À medida que os fragmentos passam pelo *laser*, seus corantes fluorescentes são ativados e a fluorescência resultante é detectada por um escâner óptico. Cada corante emite fluorescência de um comprimento de onda característico, que é lido pelo escâner óptico. A informação é alimentada em um computador para interpretação e os resultados são impressos como um conjunto de picos em um gráfico (ver Figura 19.23). As máquinas de sequenciamento automáticas podem ter 96 ou mais tubos capilares, permitindo que 50.000 a 60.000 pb de sequência sejam lidos em algumas horas.

Conceitos

O método de sequenciamento didesoxi usa ddNTPs, que terminam a síntese de DNA em bases específicas.

Checagem dos conceitos 8

Na reação de sequenciamento didesoxi, o que termina a síntese de DNA em determinada base?

a. A ausência de uma base no ddNTP paralisa a DNA polimerase.
b. O ddNTP provoca uma ruptura no esqueleto de açúcar-fosfato.
c. A DNA polimerase não incorpora um ddNTP na fita de DNA crescente.
d. A ausência de um grupo 3′-OH no ddNTP evita a adição de outro nucleotídio.

Tecnologias de sequenciamento de próxima geração

Métodos mais novos, chamados de **tecnologias de sequenciamento de próxima geração**, fazem sequenciamento centenas de vezes de modo mais rápido e menos dispendioso que o tradicional método de sequenciamento de Sanger. As tecnologias de sequenciamento de próxima geração fazem o sequenciamento em paralelo, o que significa que centenas de milhares ou até milhões de fragmentos de DNA são sequenciados simultaneamente.

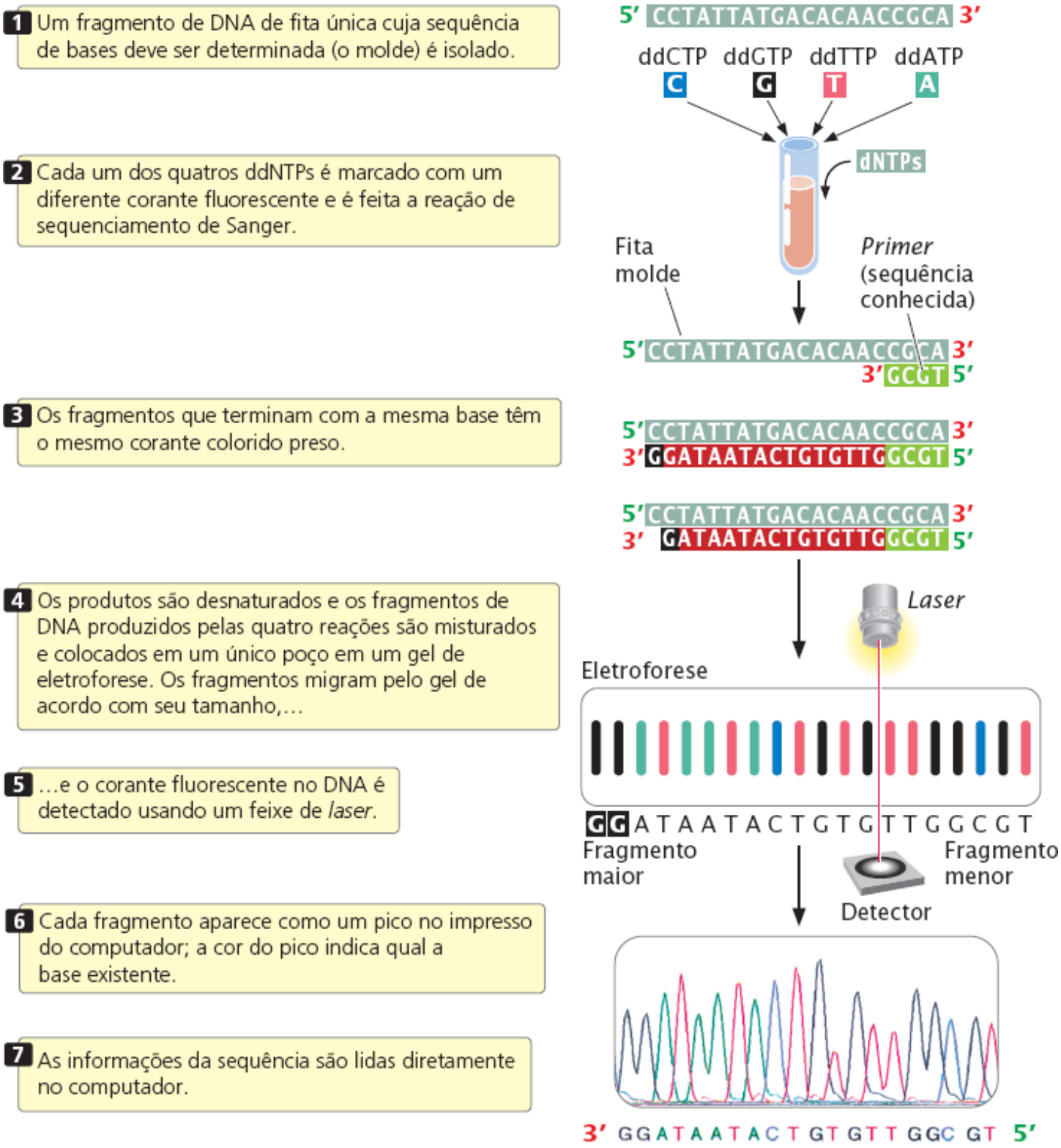


Figura 19.23 O método de sequenciamento didesoxi pode ser automatizado.

Pirossenquenciamento. Um tipo de sequenciamento de próxima geração, chamado de pirossequenciamento, baseia-se na síntese de DNA: os nucleotídios são adicionados um por vez na ordem especificada pelo DNA molde, e a adição de um nucleotídio específico é detectada por um *flash* de luz, gerado à medida que o nucleotídio é adicionado.

Para fazer o pirossequenciamento, o DNA a ser sequenciado é primeiro fragmentado (**Figura 19.24 A**). Adaptadores, compostos por um segmento curto de nucleotídios, são adicionados a cada fragmento. O adaptador fornece uma sequência conhecida para iniciar uma reação de PCR. Os fragmentos de DNA são usados para fazer uma fita única. Em uma versão do pirossequenciamento, cada fragmento é conectado a uma microesfera separada e circundado por uma gotícula de solução (**Figura 19.24 B**). A microesfera é usada para segurar o DNA e posteriormente o deposita em uma placa para a reação de sequenciamento (ver a seguir). Dentro da gotícula, o fragmento é então amplificado por PCR, e as cópias de DNA se mantêm conectadas à microesfera. Após a amplificação por PCR, as microesferas são misturadas com a DNA polimerase e depositadas em uma placa contendo mais de um milhão de poços. Cada microesfera é depositada em um diferente poço.

A reação de sequenciamento ocorre em cada poço e se baseia na síntese de DNA. Recordando o Capítulo 12, o substrato para a síntese de DNA é um trifosfato de desoxinucleosídio, composto por um açúcar desoxirribose conectado a uma base e três fosfatos. No processo de síntese de DNA, dois fosfatos (pirofosfato) são clivados e o nucleotídio resultante é conectado à extremidade 3′ da cadeia de DNA crescente. Uma solução contendo um tipo específico de trifosfato de desoxinucleosídio, por exemplo, trifosfato de desoxiadenosina, é passada através dos poços (**Figura 19.24 C**). Se o molde em determinado poço especifica um nucleotídio adenina na próxima posição na cadeia crescente, então o pirofosfato é clivado do trifosfato de nucleosídio e o nucleotídio adenina é adicionado. Uma reação química usa o pirofosfato produzido na reação para gerar um *flash* de luz, que é medido por um detector óptico. A luz emitida em cada poço é proporcional ao número de nucleotídios adicionados: se o molde em um poço especifica três nucleotídios adenina sucessivos (As), então três nucleotídios são adicionados e três vezes mais luz é emitida do que se fosse adicionado apenas um A. Se a posição no molde especifica outra base que não a adenina, nenhum nucleotídio é adicionado, não é produzido pirofosfato e não é emitida luz.

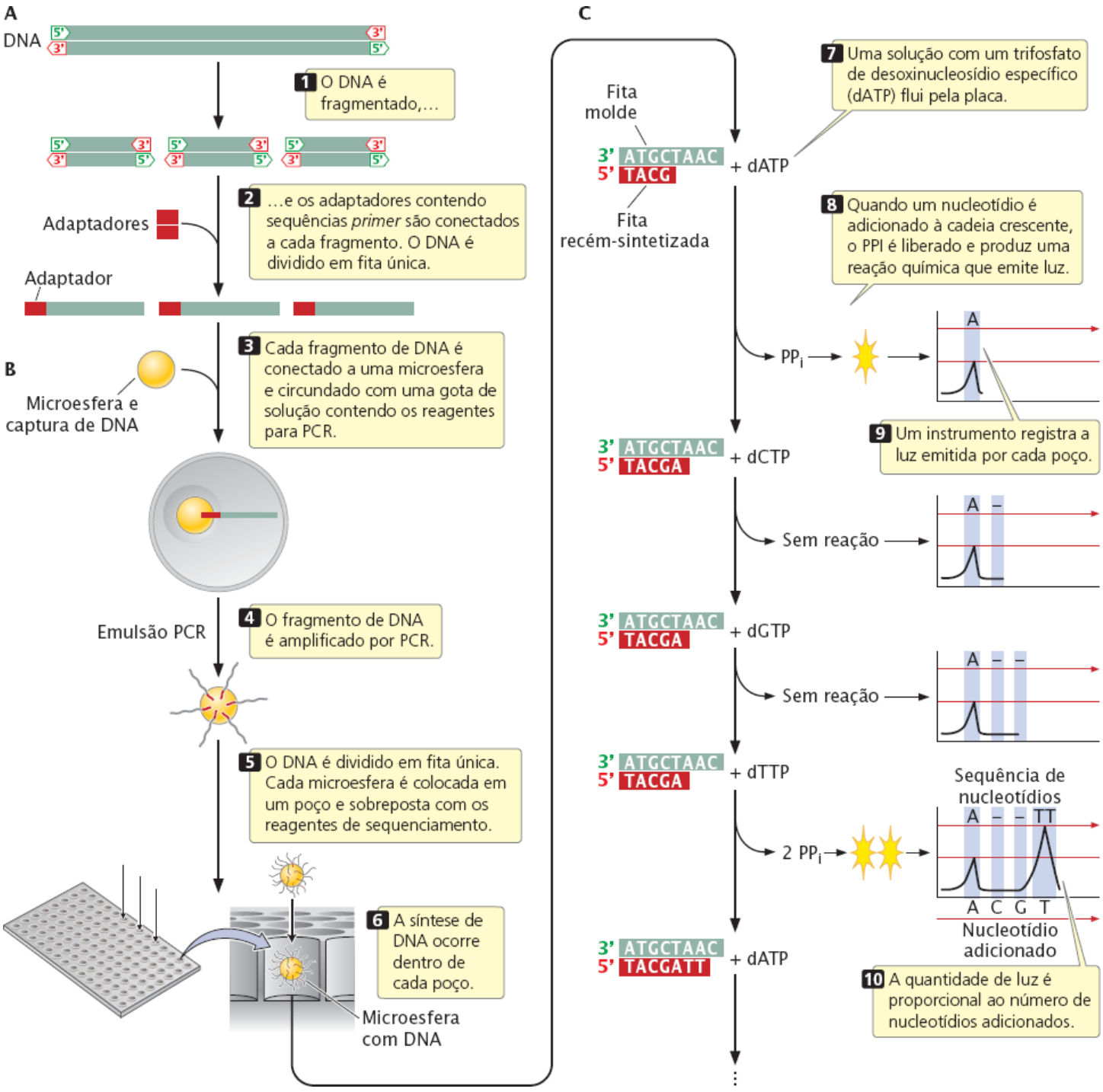


Figura 19.24 Os métodos de sequenciamento de próxima geração são capazes de determinar simultaneamente a sequência de centenas de milhares ou milhões de fragmentos de DNA. O pirossequenciamento está ilustrado nesta imagem.

Como foi mencionado, a primeira solução que passou pela placa tinha trifosfato de adenosina e permite que os nucleotídios adenina sejam adicionados ao molde. Cada poço com um molde que especifica adenina na posição seguinte gerará um *flash* de

luz. Então, uma solução com um tipo diferente de trifosfato de nucleosídio, por exemplo, trifosfato de desoxiguanosina, é passada pelos poços. Qualquer fragmento que especifique G na próxima posição na sua cadeia crescente adicionará o nucleotídio guanina e emitirá um *flash* de luz. Os trifosfatos de nucleotídios são passados pelos poços em uma ordem predeterminada e a luz emitida por cada poço é medida. Desta forma, centenas de milhares ou milhões de fragmentos de DNA são sequenciados simultaneamente com base na ordem na qual os nucleotídios são adicionados à extremidade 3′ da cadeia crescente. O pirossequenciamento determina apenas a sequência de fragmentos em cada poço; ele não permite, por si só, que as sequências destes fragmentos sejam montadas novamente na sequência do fragmento original inteiro do DNA. A remontagem dos fragmentos sequenciados em uma tira contínua de DNA é um problema comum no sequenciamento de genomas, e será discutido no Capítulo 20.

Sequenciamento illumina. São usadas várias outras formas de sequenciamento de próxima geração. O sequenciamento *illumina* emprega uma tecnologia semelhante à do método didesoxi de Sanger. São usados nucleotídios especiais que têm um marcador fluorescente preso, com um marcador colorido diferente para cada tipo de nucleotídio. Cada nucleotídio também tem um grupo químico (um terminador) que, uma vez incorporado à cadeia de DNA crescente, evita a incorporação de qualquer nucleotídio adicional. Este é semelhante ao término causado pelos didesoxinucleotídios no sequenciamento de Sanger. Entretanto, aqui, o terminador é reversível – ele pode ser removido quimicamente. Para fazer o sequenciamento, o DNA é primeiro fragmentado em milhões de fragmentos pequenos sobrepostos. Os fragmentos são presos a uma lâmina e então amplificados, criando grupos de até 1.000 cópias de cada fragmento próximos na lâmina. Os fragmentos são então desnaturados e é adicionada uma solução de *primers*, DNA polimerase e nucleotídios especiais. O *primer* se prende a cada molde de DNA e o primeiro nucleotídio é incorporado à fita recém-sintetizada. A solução é lavada e o marcador no nucleotídio incorporado é excitado com um *laser*, que gera fluorescência. Como mencionado anteriormente, cada tipo de nucleotídio (A, T, G, ou C) tem um diferente marcador fluorescente colorido, então a cor da luz produzida revela o tipo de nucleotídio adicionado. O terminador e o marcador fluorescente são removidos quimicamente e o processo é repetido. À medida que os nucleotídios são adicionados um por vez, a sequência é lida como uma série de *flashes* de luz colorida a partir de cada grupo de DNA. Centenas de milhares de grupos de DNA, cada um com cópias de um fragmento diferente de DNA, são sequenciados simultaneamente, permitindo que grandes quantidades de DNA sejam sequenciadas em um curto período de tempo.

A maioria das técnicas de sequenciamento de próxima geração lê fragmentos de DNA mais curtos que as reações de sequenciamento de Sanger, mas como centenas de milhares ou milhões de fragmentos são sequenciados simultaneamente, estes métodos são muito mais rápidos que a tecnologia de sequenciamento tradicional de Sanger.

Tecnologia de sequenciamento de terceira geração. Até métodos de sequenciamento mais avançados e rápidos, chamados de sequenciamento de terceira geração, estão em desenvolvimento. Por exemplo, o sequenciamento nanoporo determina a sequência de moléculas individuais de DNA. Neste método, uma única fita de DNA é passada por um orifício minúsculo, um nanoporo em uma membrana. À medida que a molécula passa pelo nanoporo, ela rompe uma corrente elétrica na membrana, e a natureza da ruptura é afetada pelo formato da molécula que passa pelo nanoporo. Cada uma das quatro bases de DNA provoca uma ruptura característica, então a sequência de DNA pode ser lida ao analisar a corrente da membrana à medida que a fita passa, um nucleotídio por vez, através do nanoporo. Centenas de milhares de nanoporos podem ser criados em um único *chip*, de modo que muitos fragmentos de DNA podem ser lidos simultaneamente. Um dos objetivos da tecnologia de sequenciamento de terceira geração é desenvolver um método que possa sequenciar um genoma humano inteiro por menos de mil dólares.

Conceitos

Os métodos de sequenciamento de próxima e terceira geração sequenciam muitos fragmentos de DNA simultaneamente e fornecem uma determinação mais rápida e menos onerosa de uma sequência de bases de DNA do que o método de sequenciamento de Sanger.

Fingerprinting do DNA

O uso de sequências de DNA para identificar pessoas é chamado de ***fingerprinting* do DNA** ou perfil de DNA. Como alguns pares do genoma são muito variáveis, a sequência de DNA de cada pessoa é única e, como uma impressão digital tradicional, fornece uma característica distinta que permite a identificação.

Atualmente, a maior parte do *fingerprinting* do DNA utiliza **microssatélites** ou **repetições curtas em *tandem*** (STRs), que são sequências de DNA muito curtas repetidas em *tandem* (ver *Capítulo 11*). Estas sequências repetidas são encontradas em muitos *loci* pelo genoma humano. As pessoas apresentam variação no número de cópias de sequências repetidas que elas têm

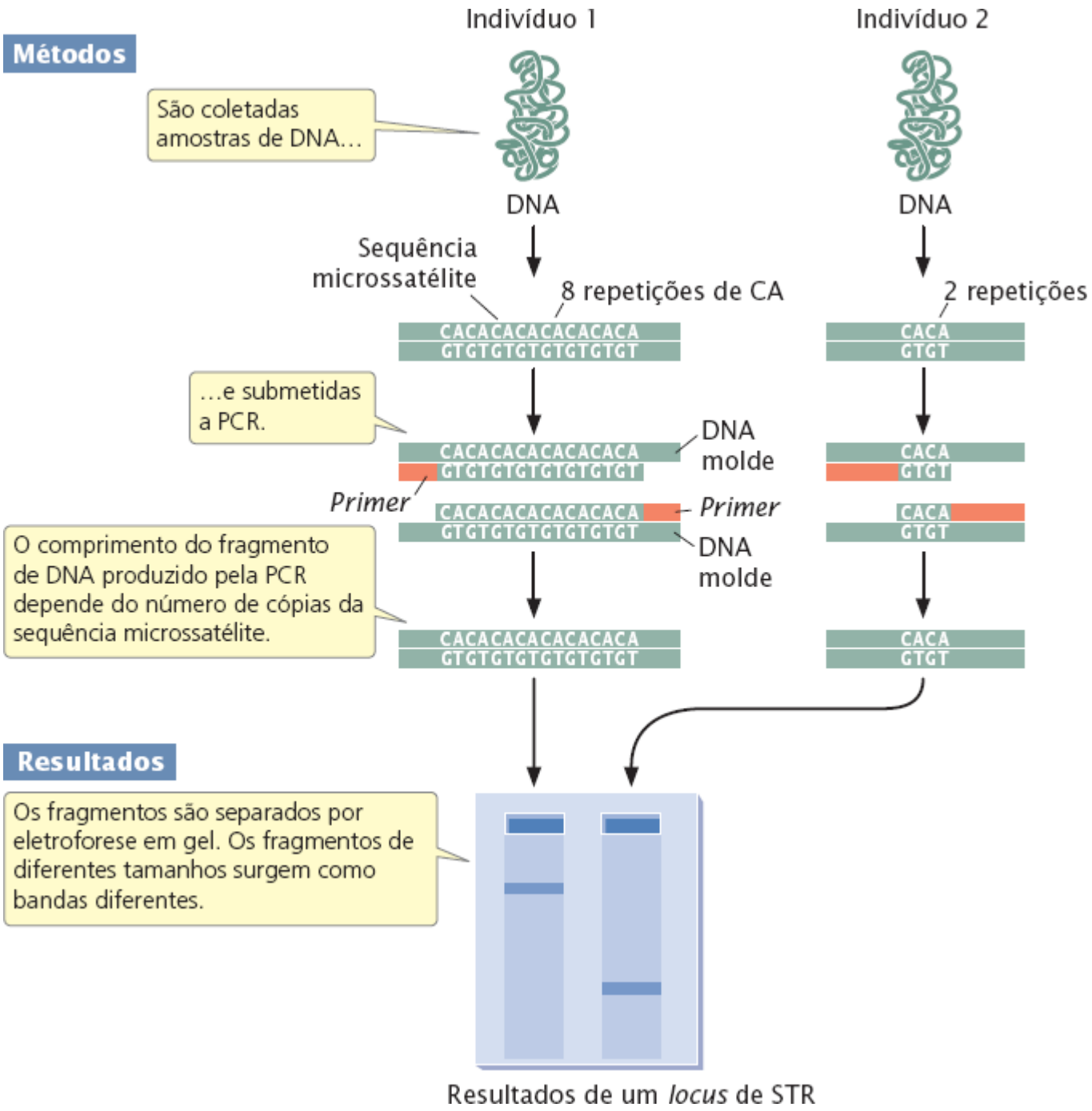
em cada *locus*. As STRs são detectadas com PCR, usando *primers* que flanqueiam as repetições microssatélites de modo que um fragmento de DNA com as sequências repetidas é amplificado (**Figura 19.25**). O comprimento do segmento amplificado depende do número de repetições; o DNA de uma pessoa com mais repetições vai produzir um segmento amplificado maior que o DNA de uma pessoa com poucas repetições.

Os *primers* usados na reação de PCR são marcados com um marcador fluorescente de modo que os fragmentos de DNA resultantes possam ser detectados com um *laser*. *Primers* de diferentes *loci* de STR são marcados com diferentes *primers* coloridos, então podem ser diferenciados produtos de diferentes tamanhos de diferentes *loci*. Após a PCR, os fragmentos são separados em um gel ou por máquina de eletroforese capilar. Na eletroforese capilar, a presença de cada fragmento é detectada à medida que ele migra pelo *laser* e um computador então calcula o tamanho de cada fragmento com base na sua taxa de migração. Os fragmentos são representados como picos em um gráfico; a distância no eixo horizontal representa o tamanho do fragmento, enquanto a altura do pico representa a quantidade de DNA (**Figura 19.26**). Os homozigotos para um alelo STR têm um único pico alto, os heterozigotos têm dois picos mais baixos. Quando são examinados vários *loci* microssatélites diferentes, a probabilidade de que duas pessoas tenham o mesmo conjunto de padrões se torna muito menor, exceto que sejam gêmeos idênticos.

O Federal Bureau of Investigation (FBI) desenvolveu um sistema que usa 13 *loci* STR (**Quadro 19.3**) que são comumente empregados para identificar pessoas e solucionar crimes. Estes *loci* constituem o Combined DNA Index System (CODIS). Cada *locus* STR no CODIS tem vários alelos e está localizado em um cromossomo humano diferente, e então a variação em cada *locus* classifica de forma independente. Quanto todos os 13 *loci* CODIS são usados juntos, a probabilidade de duas pessoas aleatoriamente selecionadas terem o mesmo perfil de DNA é menor que 1 em 10 bilhões.

Experimento

Pergunta: Como podemos identificar as pessoas com base nas diferenças no seu DNA?



Conclusão: Os padrões de fragmentos de DNA produzidos pelos indivíduos são diferentes.

Figura 19.25 O *fingerprinting* do DNA pode ser usado para identificar pessoas.

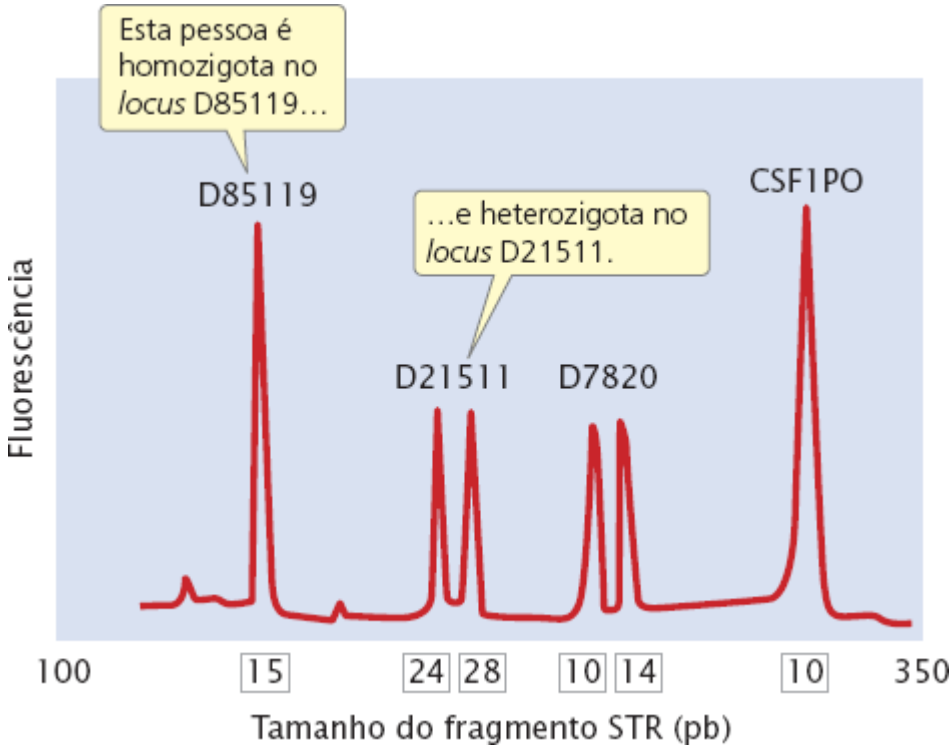


Figura 19.26 Um perfil de DNA representa o padrão dos fragmentos de DNA produzidos por PCR dos *loci* STR. Este perfil mostra os resultados a partir de quatro *loci* para STR (D8S1179, D21S11, D7S820 e CSF1 P0). O número abaixo de cada pico representa o número de repetições STR naquele fragmento de DNA.

Quadro 19.3 Características dos 13 *loci* STR usados no CODIS para *fingerprinting* do DNA.

Nome do *locus*	Cromossomo	Número de repetições	Número de alelos*
CSF1PO	5	5 a 17	10
FGA	4	12 a 51	23
TH01	11	3 a 14	8
TPOX	2	4 a 16	8
VWA	12	10 a 25	10
D3S1358	3	6 a 26	10
D5S818	5	4 a 29	9
D7S820	7	5 a 16	11
D8S1179	8	6 a 20	11
D13S317	13	5 a 17	8
D16S539	16	4 a 17	7
D18S51	18	5 a 40	19
D21S11	21	12 a 43	22

*População dos EUA.

Fonte: J. M. Butler and C. R. Hill, Forensic Science Review 24:15-26, 2012.

Comumente, o *fingerprinting* do DNA pode ser usado para confirmar que um suspeito estava presente na cena de um crime. Uma amostra de DNA do sangue, sêmen, cabelo ou outro tecido do corpo é coletada da cena do crime. Se a amostra for muito pouca, pode ser usada PCR para ampliar de modo que exista DNA suficiente para teste. Amostras adicionais de DNA são coletadas de um ou mais suspeitos. O padrão dos fragmentos de DNA produzido pelo *fingerprinting* do DNA a partir da

amostra é então comparado com os padrões produzidos pelo *fingerprinting* do DNA do suspeito. Uma combinação entre as amostras pode fornecer indício de que o suspeito estava presente na cena do crime (**Figura 19.27**).

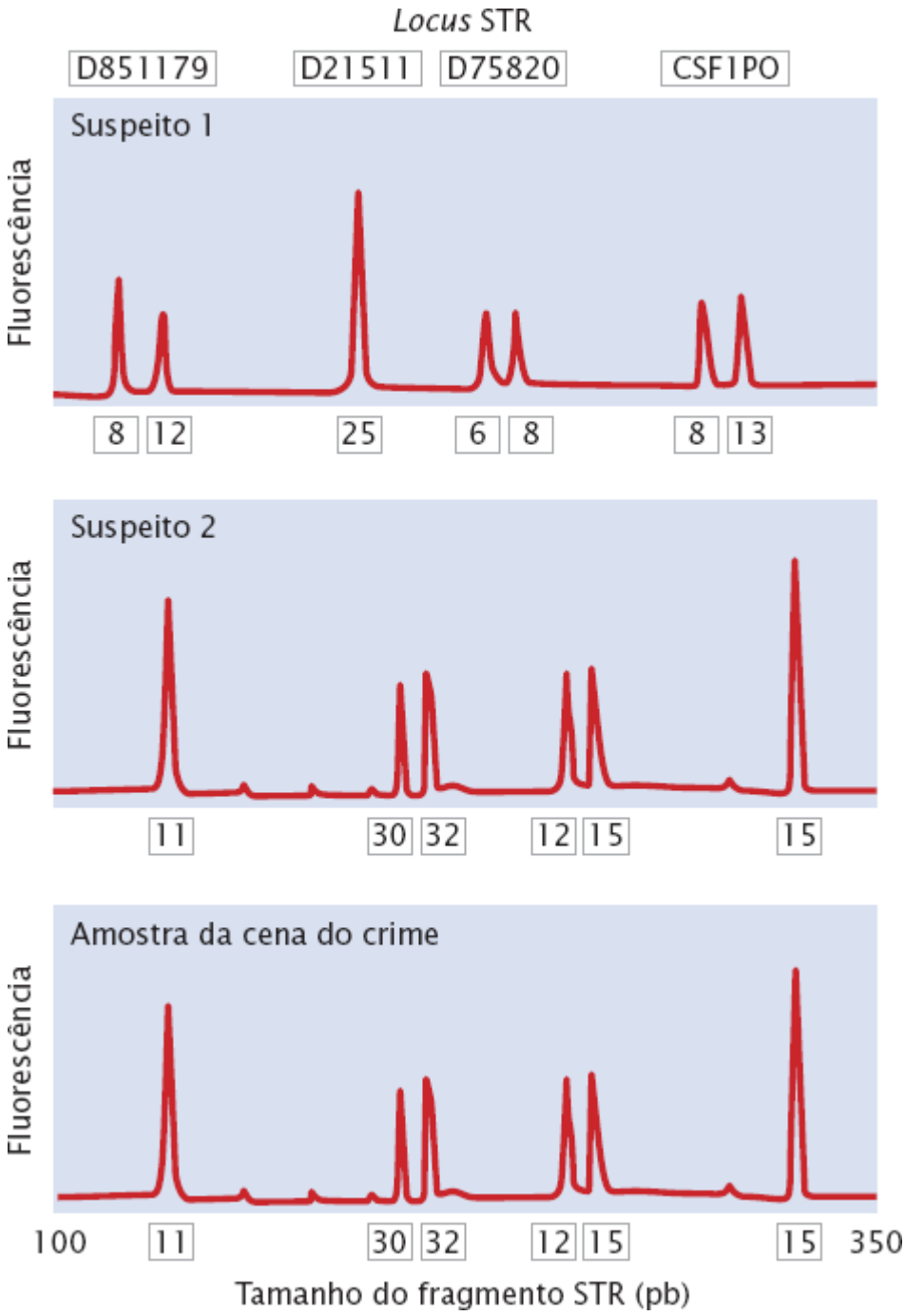


Figura 19.27 O *fingerprinting* do DNA pode ser usado para determinar a presença de um suspeito em uma cena de crime. O perfil de DNA do suspeito 2 combina com o do indício de DNA coletado na cena do crime. As imagens representam os resultados de 4 *loci* para STR.

Desde sua introdução nos anos de 1980, o *fingerprinting* do cDNA tem ajudado a condenar vários suspeitos em casos de assassinato e estupro. Os suspeitos de outros casos foram inocentados quando seu DNA não combinava com o DNA das cenas de crime. Inicialmente, calcular as chances de uma combinação (a probabilidade de que duas pessoas tenham o mesmo padrão) era controverso, e existia preocupação sobre o controle de qualidade (como uma contaminação acidental de amostras e a reprodutibilidade dos resultados) nos laboratórios onde a análise de DNA é feita. Atualmente, o *fingerprinting* do DNA se tornou uma ferramenta importante na investigação forense. Além de sua aplicação na solução de crimes, o *fingerprinting* do DNA é usado para analisar a paternidade, estudar relações genéticas entre organismos individuais em populações naturais, identificar cepas específicas de bactérias patogênicas e identificar cadáveres humanos.

Conceitos

O *fingerprinting* do DNA detecta diferenças genéticas entre pessoas ao analisar regiões muito variáveis dos

cromossomos.

Checagem dos conceitos 9

Como os microssatélites são detectados?

Aplicação | Identificação dos corpos no colapso das torres do World Trade Center

Na manhã do dia 11 de setembro de 2001, terroristas sequestraram e pilotaram dois aviões de passageiros na direção das torres gêmeas do World Trade Center na cidade de Nova York. O impacto e o posterior incêndio levaram, no decorrer de algumas horas, ao completo colapso dos 110 andares de ambas as torres, matando quase 3.000 pessoas no prédio e a equipe de resgate. A tremenda força destrutiva gerada pelo colapso das torres, com seus 325.000 metros cúbicos de concreto e 200.000 toneladas de aço, pulverizou muitos dos corpos, impossibilitando o reconhecimento deles.

Nos dias logo após o colapso do World Trade Center, os cientistas forenses começaram a tarefa de identificar as pessoas que morreram. O objetivo era obter indícios para a investigação criminal em andamento sobre o ataque e identificar os corpos para as famílias e amigos das vítimas. Essa tarefa não tinha precedente em alcance e dificuldade. Não existia uma lista completa das vítimas (como uma lista de passageiros em uma queda de avião) com a qual os investigadores pudessem comparar os cadáveres. No total, foram encontrados quase 20.000 cadáveres, variando de corpos inteiros a pequenos fragmentos de ossos queimados. Os cadáveres foram submetidos ao fogo com temperaturas superiores a 1.000°C que queimaram por mais de 3 meses. O colapso dos prédios misturou os corpos de muitas vítimas, e muitos fragmentos não foram recuperados por meses, durante os quais estiveram expostos a poeira, água, bactérias e degradação.

Os meios comuns para identificação de vítimas, itens pessoais, impressões digitais, registros dentários, tinham pouca utilidade para a maioria dos corpos do World Trade Center. A identificação da maioria dos cadáveres foi feita com o uso de *fingerprinting* do DNA.

Primeiro, o DNA foi extraído de amostras de tecido ao usar técnicas estéreis para evitar contaminação cruzada entre as amostras. Após a extração do DNA, PCR foi usada para ampliar os *loci* STR no sistema CODIS (ver seção anterior). A impressão digital do DNA gerado a partir de cada amostra de corpo foi comparada com a do DNA extraído das amostras de referência, como escovas de dente das vítimas e amostras de sangue, fornecidas pelas famílias e amigos. Se o DNA de uma parte do corpo tinha os mesmos alelos em todos os 13 *loci* como na amostra referência, era feita uma identificação positiva. Se não existisse amostra de referência, os investigadores coletavam DNA dos familiares e tentavam comparar os perfis de DNA dos corpos de seus parentes, neste caso, alguns, mas não todos os alelos STR combinariam.

Infelizmente, muitos dos corpos estavam tão degradados que sobrou pouco DNA e um ou mais dos *loci* STR não puderam ser amplificados. Para estes corpos, também foi feito o *fingerprinting* do DNA no DNA mitocondrial (ver *Capítulo 11*). Como existem muitas mitocôndrias por célula e cada mitocôndria tem múltiplas moléculas de DNA, existem mais cópias de DNA mitocondrial por célula do que DNA nuclear; o DNA mitocondrial foi extraído com sucesso e analisado a partir dos corpos de ancestrais, como os neandertais (ver *Capítulo 10*). Sozinho, o *fingerprinting* do DNA mitocondrial não foi suficiente para comprovar a identificação com alto grau de confiança (porque não existem tantas sequências que variem entre as pessoas como nos *loci* CODIS), mas, quando usado junto com os dados de pelo menos alguns *loci* STR, era possível fazer a identificação positiva.

Combinadas, estas técnicas permitiram que os restos mortais de muitas vítimas fossem identificados. Entretanto, apesar dos esforços heroicos de centenas de geneticistas moleculares, antropólogos forenses e médicos, não foram recuperados restos mortais positivamente identificados de quase metade das pessoas que se acredita que tenham morrido no desastre.

19.5 Técnicas moleculares são usadas para analisar a função dos genes

Nas seções anteriores, aprendemos sobre as poderosas técnicas moleculares para isolar, recombinar e analisar sequências de DNA. Embora elas forneçam muitas informações sobre a organização e a natureza das sequências gênicas, o objetivo principal de muitos estudos moleculares é compreender melhor a função das sequências. Nesta seção, vamos explorar algumas técnicas moleculares avançadas que são frequentemente usadas para determinar a função gênica e para compreender melhor os processos genéticos que estas sequências sofrem.

Genética direta e reversa

A abordagem tradicional para o estudo da função gênica começa com o isolamento de organismos mutantes. Por exemplo, suponha que um geneticista esteja interessado nos genes que afetam a função cardíaca nos mamíferos. Uma primeira etapa seria

encontrar indivíduos, talvez camundongos, que tenham defeitos hereditários na função cardíaca. As mutações responsáveis pelos problemas cardíacos nos camundongos poderiam ser mapeadas e os genes envolvidos poderiam ser isolados e sequenciados. As proteínas produzidas pelos genes poderiam ser previstas a partir das sequências de genes e isoladas. Finalmente, a bioquímica das proteínas poderia ser estudada e seu papel na função cardíaca compreendido. Esta abordagem, que começa com um fenótipo (um mutante) e continua em um gene que codifica o fenótipo, é chamada de **genética direta**.

Uma abordagem alternativa é começar com um genótipo, uma sequência de DNA, e continuar com o fenótipo ao alterar a sequência ou inibir sua expressão. Um geneticista pode começar com um gene de função desconhecida, induzir mutações nele e então observar que efeito estas mutações têm no fenótipo do organismo. Esta abordagem é chamada de **genética reversa**. Atualmente, ambas as abordagens de genética direta e reversa são muito usadas para análises de função dos genes.

Criação de mutações aleatórias

A genética direta depende da identificação e do isolamento de mutações aleatórias que afetem um fenótipo de interesse. No início do estudo da genética, os geneticistas eram forçados a confiar em mutações que surgiam naturalmente, que eram raras e podiam ser detectadas apenas se fossem examinados muitos organismos. A descoberta de agentes mutagênicos, fatores ambientais que aumentam a taxa de mutação (ver *Capítulo 18*), forneceu meios para aumentar o número de mutações nas populações experimentais de organismos. Um dos primeiros exemplos de mutações experimentalmente criadas foi o uso de raios X por Hermann Muller em 1927 para induzir mutações ligadas ao X na *Drosophila melanogaster*.

Atualmente, usam-se radiação, mutágenos químicos e elementos de transposição para criar mutações para análise genética. Para determinar todos os genes que poderiam afetar um fenótipo, é desejável criar mutações em tantos genes quanto possível, ou seja, saturar o genoma com mutações. Esse procedimento é feito com uma triagem mutagênica, que será descrita no Capítulo 20.

Mutagênese direcionada pelo sítio

A genética reversa depende da capacidade de criar mutações, não aleatoriamente, mas em sequências de DNA específicas e depois estudar os efeitos destas mutações no organismo. As mutações são induzidas em locais específicos por meio de um processo chamado de **mutagênese direcionada pelo sítio**.

Já foram desenvolvidas várias estratégias diferentes para esta mutagênese. Uma estratégia que é usada nas bactérias é cortar uma sequência curta de nucleotídios com enzimas de restrição e substituir com um oligonucleotídio curto sintético que tenha a sequência mutante desejada. O sucesso deste método depende da disponibilidade dos sítios de restrição flanqueadores da sequência a ser alterada.

Pode ser usada a **mutagênese direcionada por oligonucleotídio** se não estiverem disponíveis sítios de restrição adequados (**Figura 19.28**). Neste método, é produzido um oligonucleotídio de fita única diferente da sequência-alvo em uma ou algumas bases. Como eles diferem em apenas algumas bases, o DNA-alvo e o oligonucleotídio vão parear. Ao parear com o DNA-alvo, o oligonucleotídio pode agir como *primer* ao iniciar a síntese de DNA, que produz uma molécula de fita dupla com um pareamento errado na região do *primer*. Quando este DNA é transferido para as células bacterianas, as bases com pareamento errado serão reparadas pelas enzimas bacterianas. Cerca de metade do tempo as bases normais serão trocadas por bases mutantes e cerca de metade do tempo as bases mutantes serão trocadas por bases normais. As bactérias são então rastreadas à procura do gene mutante.

Conceitos

A genética direta começa com um fenótipo e detecta e analisa o genótipo que provoca o fenótipo. A genética reversa começa com uma sequência de genes e por meio de análise, determina o fenótipo que ela codifica. Mutações específicas podem ser introduzidas em sítios específicos em um gene por meio de mutagênese direcionada pelo sítio ou por oligonucleotídio.

Checagem dos conceitos 10

Um geneticista interessado na função imunológica induz mutações aleatórias em vários genes específicos nos camundongos e então determina qual dos camundongos mutantes tem a função imunológica comprometida. Este procedimento é um exemplo de:

a. Genética direta.
b. Genética direta e reversa.
c. Genética reversa.
d. Nem genética direta nem reversa.

Animais transgênicos

Outra forma que pode ser usada para analisar a função do gene é adicionar sequências de DNA de interesse ao genoma de um organismo que normalmente não tem tais sequências e então observar o efeito que a sequência introduzida tem sobre o fenótipo do organismo. Este método é uma forma de genética reversa. Um organismo modificado de forma permanente pela adição de uma sequência de DNA ao seu genoma é chamado de *transgênico*, e o DNA estranho que ele carreia é chamado de **transgene**. Aqui, vamos analisar as técnicas para a criação de camundongos transgênicos, que são usados no estudo da função dos genes humanos porque eles podem ser geneticamente manipulados por meios que são impossíveis nos humanos e, como mamíferos, eles são mais semelhantes aos humanos que as moscas-da-fruta e outros organismos-modelo da genética.

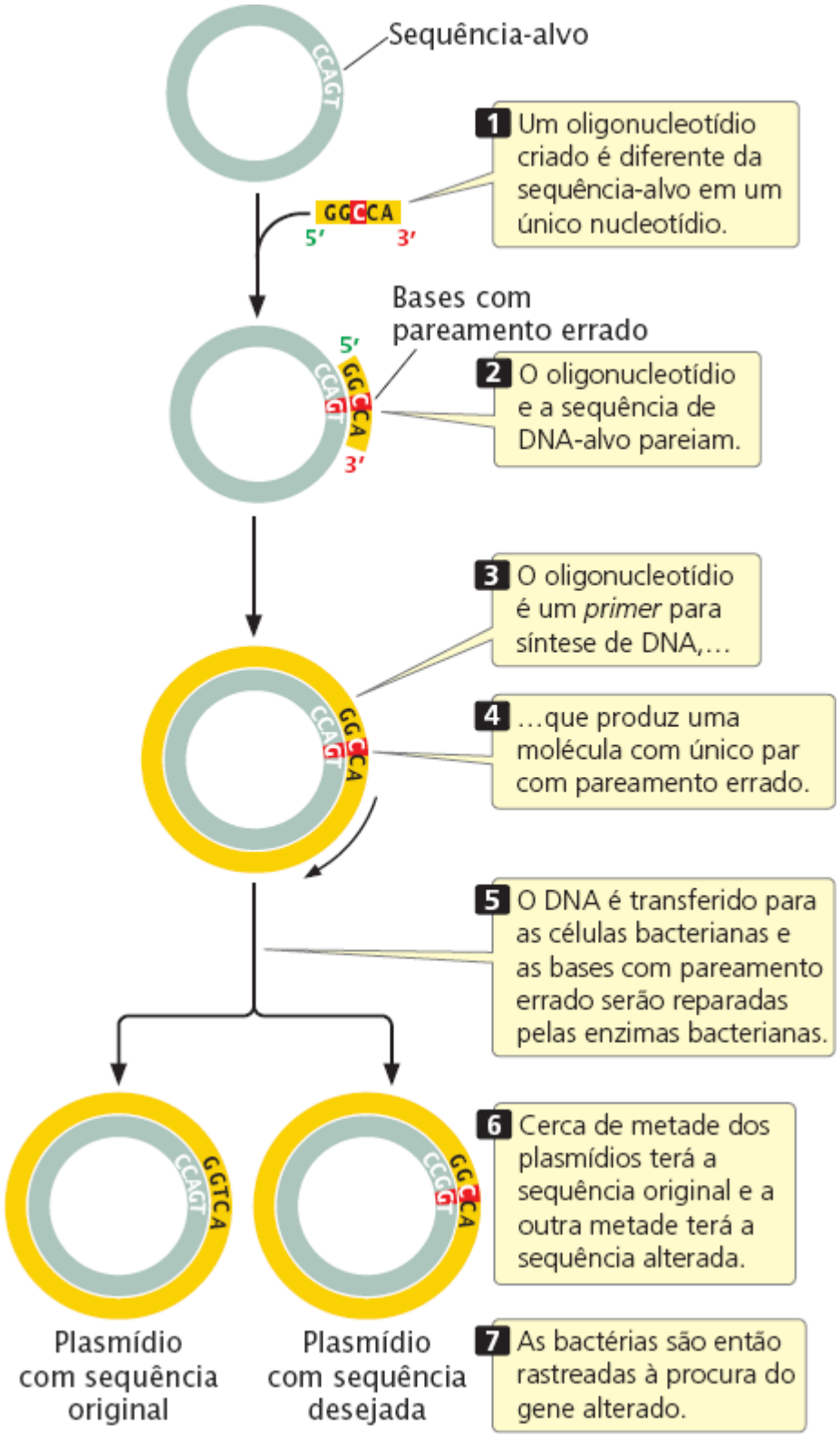
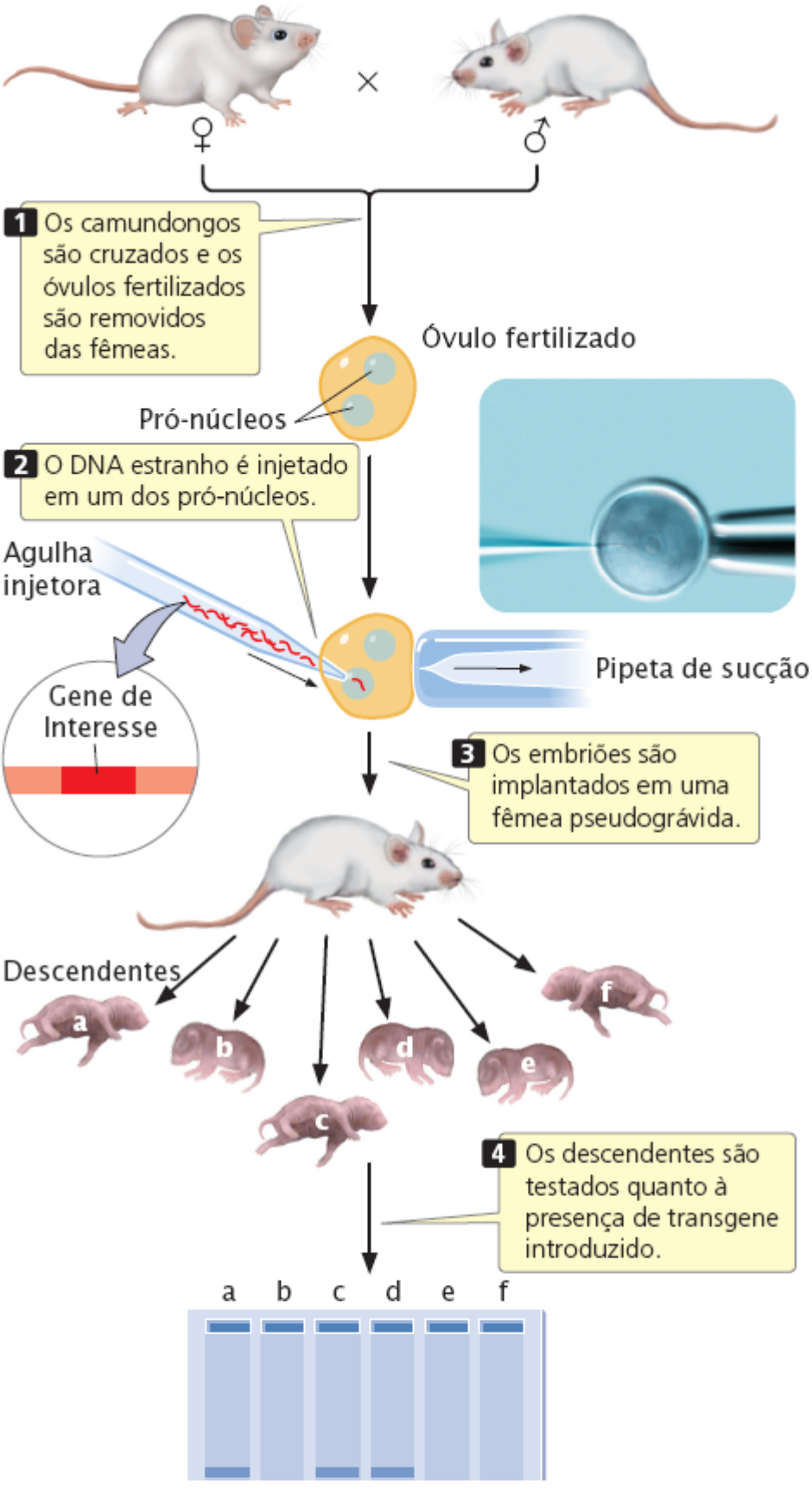


Figura 19.28 A mutagênese direcionada por oligonucleotídio é usada para estudar a função do gene quando os sítios de restrição adequados não estão disponíveis.

Os ovócitos dos camundongos e outros mamíferos são grandes o suficiente para que o DNA seja injetado neles diretamente. Imediatamente após a penetração de um espermatozoide, um óvulo de camundongo fertilizado tem dois pró-núcleos, um do espermatozoide e um do óvulo; eles vão se fundir posteriormente para formar o núcleo do embrião. Dispositivos mecânicos podem manipular agulhas de vidro extremamente finas e ocas para injetar o DNA diretamente em um dos pró-núcleos de um óvulo fertilizado (**Figura 19.29**). Algumas cópias de DNA clonado linear são injetadas em um pró-núcleo e em alguns dos óvulos fertilizados, as cópias do DNA clonado se integram de forma aleatória a um dos cromossomos mediante um processo chamado de recombinação não homóloga. Após a injeção, os embriões são implantados em uma fêmea pseudográvida, uma mãe de aluguel que foi fisiologicamente preparada para gestação ao cruzar com um macho vasectomizado.



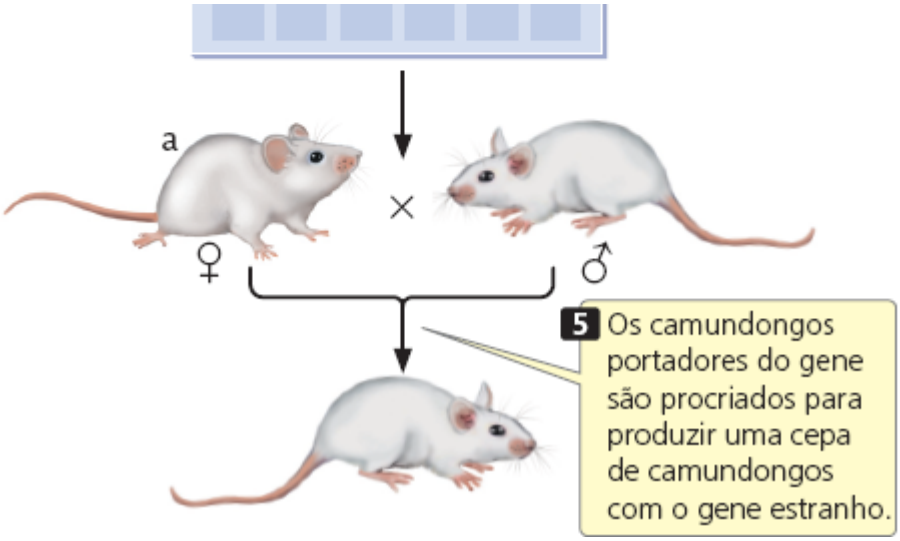


Figura 19.29 Os animais transgênicos têm genomas que são alterados de forma permanente por meio da tecnologia do DNA recombinante. Na fotografia, um embrião de camundongo recebe injeção de DNA. (Fotografia: Chad Davis/PhotoDisc.)

Cerca de apenas 10 a 30% dos embriões sobrevivem e, dos que sobrevivem, apenas alguns têm uma cópia do DNA clonado integrado de forma estável em um cromossomo. Contudo, se várias centenas de embriões recebem a injeção e são implantados, existe uma boa chance de que nasça um ou mais camundongos cujos cromossomos têm o DNA exógeno. Além disso, como o DNA foi injetado no estágio unicelular do embrião, estes camundongos carreiam o DNA clonado em todas as células dos seus corpos, incluindo as células reprodutivas e, portanto, passam o DNA exógeno para seus descendentes. Por meio do intercruzamento, pode ser criada uma cepa de camundongos que carreiem o gene exógeno.

Foi provado que os camundongos transgênicos são úteis no estudo da função dos genes. Por exemplo, foi obtida a prova de que o gene *SRY* (ver *Capítulo 4*) é o gene determinante dos machos nos camundongos ao injetar uma cópia do gene *SRY* embriões XX e observar que estes camundongos se desenvolveram como machos. Além disso, os pesquisadores criaram várias cepas de camundongos transgênicos que servem como modelos experimentais para doenças genéticas humanas.

Camundongos knockout

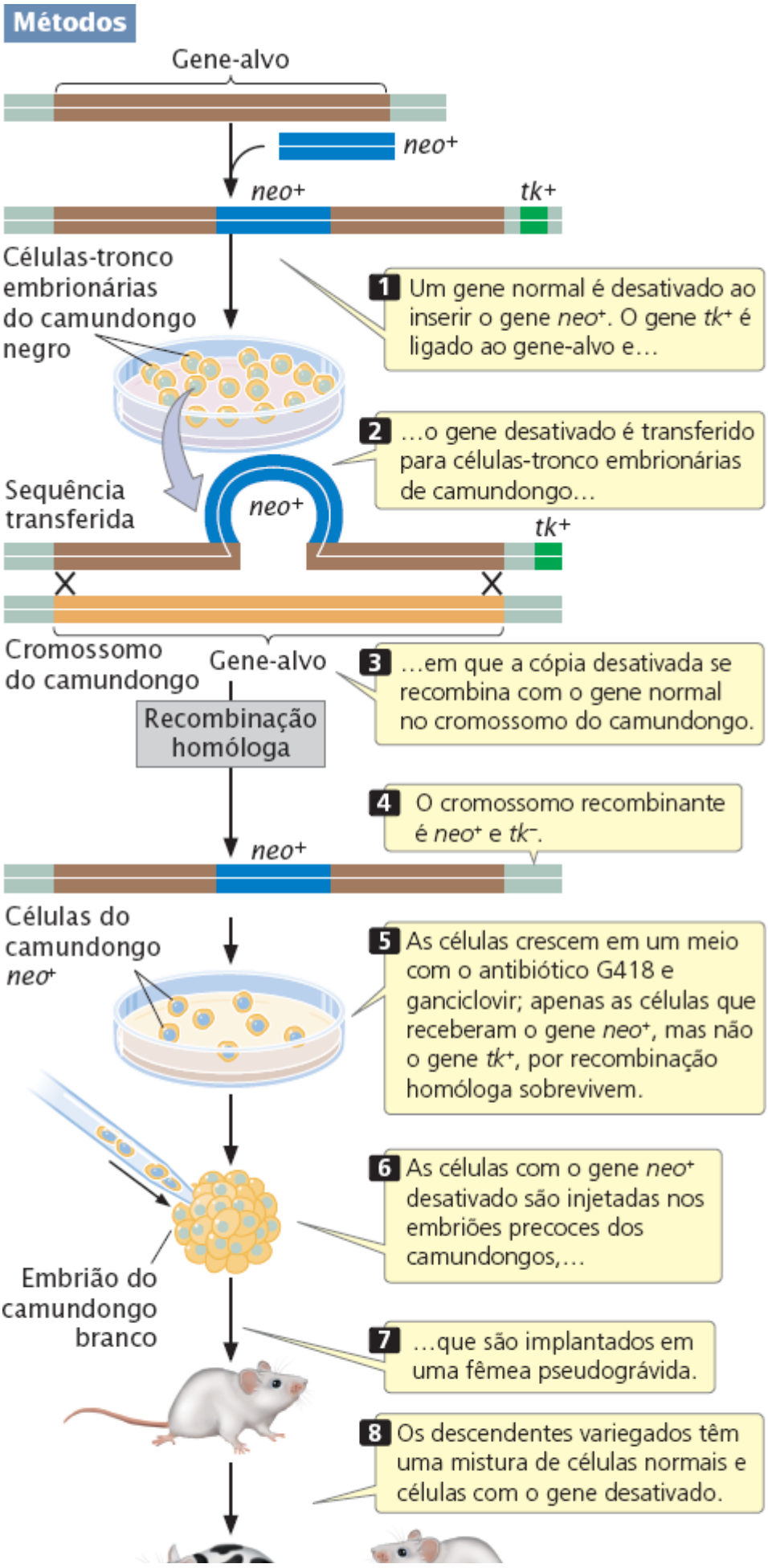
Uma variante útil da abordagem transgênica é produzir camundongos nos quais um gene normal não sofreu apenas uma mutação, mas foi totalmente desativado. Estes animais, chamados de **camundongos *knockout***, são particularmente úteis para determinar a função de um gene: um genótipo do camundongo *knockout* fornece uma boa indicação da função do gene ausente.

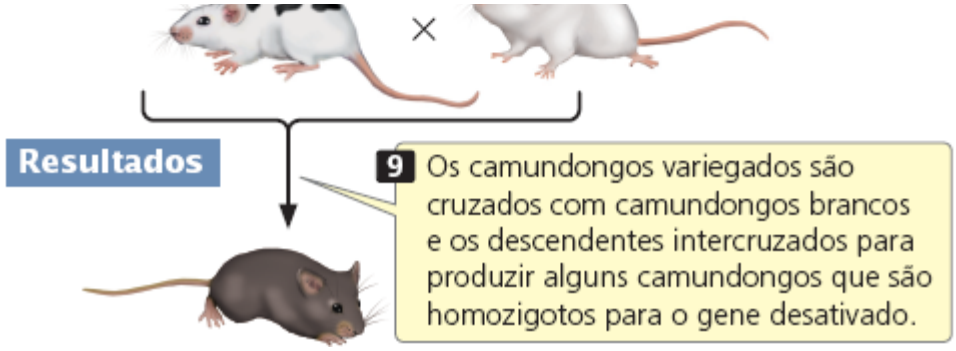
A criação dos camundongos *knockout* começou quando um gene normal é clonado nas bactérias e então “nocauteado” ou desativado. Existem várias formas para desabilitar um gene, mas um método mais comum é inserir um gene chamado *neo*, que confere resistência ao antibiótico G418 no meio do gene-alvo (**Figura 19.30**). A inserção do *neo* interrompe o gene-alvo e fornece um marcador conveniente para encontrar cópias do gene desativado. Além disso, um segundo gene, em geral o gene da timidinoquinase do herpes-vírus simples (*tk*), está ligado ao gene dividido. O gene desativado é então transferido para células embrionárias de camundongos, onde podem trocar lugares com a cópia normal no cromossomo do camundongo graças à recombinação homóloga.

Após o gene desativado ser transferido para as células embrionárias, as células são rastreadas por meio de adição do antibiótico G418 ao meio. Apenas as células com o gene desativado com o gene *neo* inserido sobreviverão. Como a frequência da recombinação não homóloga é maior que a da recombinação homóloga e o gene-alvo intacto é substituído pela cópia desativada apenas por meio da recombinação homóloga, é necessário um meio para selecionar os recombinantes homólogos raros. O gene *tk* viral torna a célula sensível ao ganciclovir. Assim, as células transferidas que crescem no meio com G418 e ganciclovir terão o gene *neo* (gene-alvo desativado), mas não o gene *tk* adjacente, porque este gene será eliminado no evento de *crossing over* duplo. Estas células têm os recombinantes homólogos desejados. Os recombinantes não homólogos (inserções aleatórias) terão os genes *neo* e *tk*, e estas células transferidas morrerão no meio selecionado por causa do ganciclovir. As células sobreviventes são injetadas em um embrião de camundongo em estágio inicial que é implantado em uma fêmea pseudográvida de camundongo. As células no embrião com o gene desativado e as células embrionárias normais com o gene do tipo selvagem se desenvolverão juntas, produzindo uma quimera, um camundongo que é uma mistura genética de dois tipos de células. A produção de camundongos quimera não é por si só desejável, mas é impossível substituir todas as células do embrião com as células injetadas.

Experimento

Pergunta: Como a função de um gene pode ser determinada?





Conclusão: Ao usar este processo, são produzidos camundongos sem a cópia funcional do gene, ou seja, o gene está desativado. O fenótipo dos camundongos knockout revela a função do gene.

Figura 19.30 Os camundongos *knockout* têm um genoma no qual um gene foi desativado.

Os camundongos quimeras podem ser identificados facilmente se as células embrionárias injetadas vierem de um camundongo negro e os embriões em que foram injetadas vierem de um camundongo branco; as quimeras resultantes terão o pelo variado de preto e branco. Algumas das quimeras têm o gene *knockout* nas suas células de linhagem germinativa e podem transmiti-lo para a próxima geração. As quimeras são cruzadas com os camundongos brancos; qualquer descendente negro é heterozigoto para o *knockout*. Os descendentes negros podem ser intercruzados para produzir alguns descendentes que sejam homozigotos para o gene *knockout*. Os efeitos de desabilitar um gene particular podem ser observados nestes camundongos homozigotos. Três cientistas que ajudaram a desenvolver as técnicas para criar os camundongos *knockout*, Mario Capecchi, Oliver Smithies e Martin Evans, receberam o Prêmio Nobel em Fisiologia e Medicina em 2007 por seu trabalho.

Foram desenvolvidas técnicas adicionais para nocautear genes apenas em alguns tecidos ou em alguns momentos; são chamadas de *knockout* adicionais. Uma variação do procedimento *knockout* é inserir nos camundongos uma sequência particular de DNA em uma localização cromossômica conhecida. Por exemplo, os pesquisadores podem inserir a sequência do alelo responsável por uma doença humana no mesmo *locus* nos camundongos, criando um modelo preciso de camundongo da doença humana. Os camundongos que carreiam sequências inseridas em localizações específicas são chamados de **camundongos *knock-in***.

Conceitos

Um camundongo transgênico é produzido pela injeção de DNA clonado no pró-núcleo de um óvulo fertilizado, seguida pela implantação do óvulo em um camundongo fêmea. Nos camundongos *knockout*, o DNA injetado tem uma mutação que desabilita um gene. Dentro do embrião do camundongo, a cópia desativada do gene pode trocar com a cópia normal do gene por meio da recombinação homóloga.

Checagem dos conceitos 11

Qual é a vantagem de usar o gene *neo* para interromper a função de um gene nos camundongos *knockout*?

a. O gene *neo* produz um antibiótico que mata as células indesejáveis.
b. O gene *neo* está no lado certo para desabilitar outros genes.
c. O gene *neo* fornece um marcador selecionável para encontrar as células que têm o gene desativado.
d. O gene *neo* produz uma toxina que inibe a transcrição do gene-alvo.

Como silenciar os genes com RNAi

Nas seções anteriores, analisamos a função dos genes ao introduzir mutações ou novas sequências de DNA no genoma e analisando o fenótipo resultante para obter informações sobre a função do DNA alterado ou introduzido. Poderíamos analisar a função dos genes ao desligar temporariamente um gene e observar que efeito a ausência do produto do gene teria sobre o fenótipo. Por muitos anos, não existia método para afetar de forma seletiva a expressão gênica. Entretanto, a descoberta de siRNA (pequenos RNAs de interferência) e miRNA (microRNAs; ver *Capítulos 14* e *17*) forneceu ferramentas poderosas para controlar a expressão dos genes individuais.

Lembre-se de que os siRNAs e miRNAs são pequenas moléculas de RNA que combinam com proteínas para formar o complexo silenciador induzido por RNA (RISC). Em um processo chamado de interferência por RNA ou RNAi, o RISC pareia com sequências complementares no mRNA e rompe o mRNA ou evita que ele seja traduzido. Os geneticistas moleculares

exploram esse maquinário natural para desligar a expressão de genes específicos. O estudo do efeito de silenciar um gene com siRNA pode ser uma fonte de compreensão para a função do gene.

A primeira etapa ao usar a tecnologia RNAi é projetar os siRNAs de modo que eles sejam reconhecidos e clivados pela Dicer (a proteína que processa os siRNAs, ver *Capítulo 14*). A sequência complementar deve ser única ao mRNA-alvo e não encontrada em outros mRNAs de modo que o siRNA não iniba outros mRNAs além do alvo. Programas de computador são usados para projetar siRNAs ideais.

Após uma sequência de siRNA ser projetada, ela deve ser sintetizada. Uma forma para criar um siRNA é usar um sintetizador de oligonucleotídios para sintetizar um fragmento de DNA correspondente à sequência de siRNA. O oligonucleotídio sintetizado pode ser clonado em um vetor de expressão de plasmídio entre dois promotores fortes (**Figura 19.31**). A *Escherichia coli* é transformada com o plasmídio. Dentro das bactérias, a transcrição a partir dos dois promotores seguirá em ambas as direções, produzindo duas moléculas de RNA complementares que vão parear para formar uma molécula de RNA de fita dupla reconhecida pela Dicer. Por outro lado, as sequências de DNA de fita dupla podem ser sintetizadas diretamente com um sintetizador de genes.

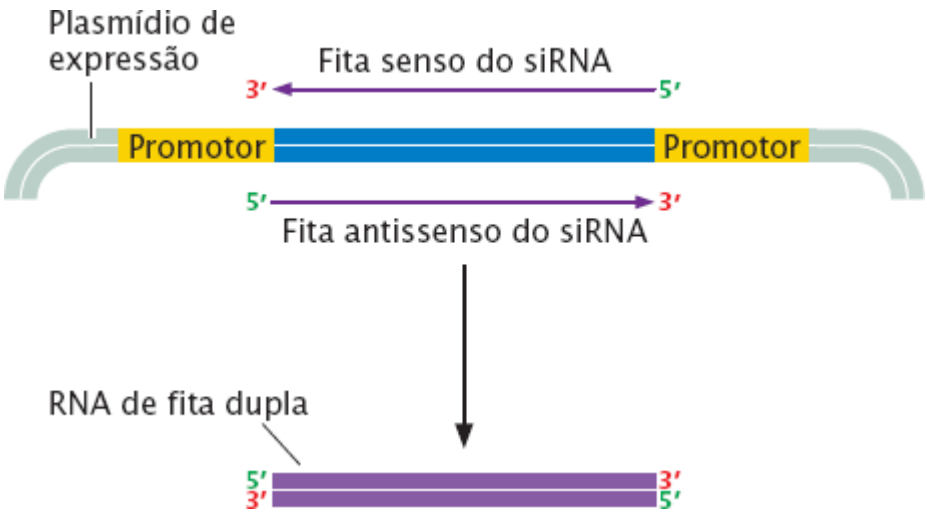


Figura 19.31 Pequenos RNAs de interferência podem ser produzidos pela clonagem de sequências de DNA correspondentes aos siRNAs entre dois promotores fortes. Quando clonadas em um vetor de expressão, ambas as fitas de DNA serão transcritas e as moléculas de RNA complementares vão dobrar para formar o RNA de fita dupla que será processado em siRNA por Dicer.

A próxima etapa é entregar o siRNA de fita dupla às células. A entrega pode ser feita de várias formas, dependendo do tipo de célula. O organismo-modelo da genética *Caenorhabditis elegans* (ver *Guia de Referência para Organismos-Modelo da Genética*) pode ser alimentado pela *E. coli* (seu alimento natural) contendo o vetor de expressão. A transcrição dentro das bactérias produz RNA de fita dupla, que os vermes ingerem e incorporam em suas células. Por outro lado, o siRNA de fita dupla pode ser injetado diretamente nas células ou na cavidade corporal. Outra abordagem é sintetizar uma sequência curta de DNA que tem complementaridade interna de modo que, quando transcrita, se dobre em um RNA grampo curto (shRNA) com uma seção de fita dupla. Dentro de uma célula os shRNAs são processados pela Dicer para produzir siRNAs responsáveis pelo silenciamento dos genes. As sequências de DNA com siRNAs podem ser introduzidas em um vetor com o uso de técnicas padrão de clonagem e o vetor pode ser usado para liberar o DNA na célula. Uma vantagem desta abordagem é que, além da adição de uma sequência de DNA, a sequência RNAi tem o potencial de se tornar uma parte permanente do genoma celular e ser transmitida para os descendentes.

Uma das principais vantagens do siRNA para controlar a expressão gênica é que ele atua em trans, ou seja, uma única cópia de um gene do siRNA desligará a expressão de ambas cópias do gene-alvo. Outra vantagem é que o gene-alvo permanece intacto e desta forma, os efeitos do silenciamento são reversíveis. Resolva o Problema 43

Aplicação | Uso do RNAi para o tratamento de doenças humanas

Além da sua importância para determinar a função do gene, o RNAi tem o potencial como agente terapêutico para o tratamento futuro de doenças humanas. Este potencial inclui usar siRNAs contra vírus RNA como o HIV, assim como usá-los para tratar doenças genéticas e câncer.

Tratamento do colesterol elevado com RNAi. A pesquisa examinou o potencial de RNAi para o tratamento dos distúrbios metabólicos do colesterol. Embora o colesterol seja essencial para vida, colesterol em excesso não é saudável: o colesterol elevado no sangue é o principal contribuinte para cardiopatia, a causa principal de morte nos EUA. O colesterol é

transportado em todo o corpo na forma de pequenas partículas chamadas de lipoproteínas, que são compostas de um cerne de lipídios circundado por uma casca de fosfolipídios e proteínas (ver Figura 6.6). A proteína ApoB é uma parte essencial das lipoproteínas. Algumas pessoas têm mutações genéticas que produzem níveis elevados de ApoB, o que predispõe para a doença coronariana. Os achados a partir de estudos sugerem que diminuir a quantidade de ApoB pode reduzir o número de lipoproteínas e baixar o colesterol sanguíneo nestas pessoas, assim como nas pessoas que têm colesterol elevado por outros motivos.

Em 2006, Tracy Zimmermann e seus colegas nas empresas Alnylam Pharmaceuticals e Protiva Biotherapeutic demonstraram que o RNAi poderia ser usado para reduzir os níveis de ApoB e o colesterol sanguíneo em primatas não humanos. Primeiro, os investigadores criaram siRNAs que miravam a expressão do gene *apoB* (apoB-siRNAs) com base na sequência conhecida do gene. Os apoB-siRNAs eram sintetizados no laboratório e compostos por 21 nucleotídios na fita senso (a fita complementar ao apoB mRNA) e 23 nucleotídios na fita complementar antissenso, com uma extremidade protuberante de dois nucleotídios.

A próxima tarefa era inserir o apoB-siRNA na célula. Embora os siRNAs sejam captados rapidamente pelas células dos invertebrados como *C. elegans*, a maioria dos siNRAs não passa com facilidade pelas membranas das células de mamíferos de forma que sejam efetivos no silenciamento gênico. Além disso, os siRNAs são rapidamente removidos de circulação. Para superar estes problemas Zimmermann e seus colegas encapsularam os apoB-siRNAs em lipídios, criando partículas estáveis de ácido nucleico-lipídio (SNALPs). Os SNALPs aumentavam muito o tempo gasto pelos siRNAs na circulação e acentuavam sua captação pela célula.

Os pesquisadores testaram os efeitos de apoB-siRNAs na redução da síntese da proteína ApoB e nos níveis de colesterol. Eles injetaram SNALPs com apoB-siRNA em duas doses nos macacos cinomolgo (**Figura 19.32 A**): 1 mg/kg e 2,5 mg/kg. Além disso, eles injetaram uma salina como controle em um terceiro grupo de macacos. E descobriram que a apoB-siRNA silenciava o gene *apoB*: 48 h após o tratamento, o mRNA de *apoB* no fígado estava reduzido em 68% para os macacos que receberam a dose de 1 mg e 90% para os que receberam a dose de 2,5 mg.

Quando os pesquisadores examinaram os níveis de colesterol sérico nos macacos, descobriram que os macacos que receberam apoB-siRNA tinham uma redução significativa nos níveis de colesterol no sangue (**Figura 19.32 B**). E o mais importante, não foram observados efeitos negativos do tratamento com siRNA. Embora seja preliminar, este estudo sugere que os siRNAs têm potencial para o futuro tratamento de doenças nos humanos.

19.6 A biotecnologia utiliza o poder da genética molecular

Além de fornecer novas informações valiosas sobre a natureza e função dos genes, as técnicas de genética molecular têm muitas aplicações práticas, incluindo a produção de produtos farmacêuticos e outras substâncias químicas, bactérias especializadas, importantes plantas para a agricultura e animais de fazenda geneticamente modificados. A tecnologia também é muito usada em exames médicos e em alguns casos é até usada para corrigir defeitos genéticos em humanos. Centenas de empresas agora especializadas em desenvolver produtos através da engenharia genética e várias corporações multinacionais investiram muito dinheiro na pesquisa da genética molecular. Como já foi discutido, a análise do DNA também é usada em investigações criminais e para a identificação de cadáveres humanos.

Produtos farmacêuticos

Os primeiros produtos comerciais a serem desenvolvidos com o uso de engenharia genética foram os produtos farmacêuticos usados no tratamento de doenças e distúrbios humanos. Em 1979, a corporação Eli Lilly começou a vender insulina humana produzida com o uso da tecnologia do DNA recombinante. O gene para a insulina humana era inserido nos plasmídios e transferido para as bactérias que produziam a insulina humana. Antes, a insulina era isolada do pâncreas de porco e vaca e alguns diabéticos desenvolviam reações alérgicas a esta proteína estranha. A insulina recombinante tem a vantagem de ser a mesma que a produzida no corpo humano. Outros produtos farmacêuticos produzidos pela tecnologia do DNA recombinante incluem o hormônio do crescimento humano (para crianças com deficiência de crescimento), fatores de coagulação (para hemofílicos) e ativador do plasminogênio tecidual (usado para dissolver coágulos de sangue em pacientes durante ataque cardíaco).

Bactérias especializadas

As bactérias participam de muitos processos industriais, incluindo a produção de etanol a partir de material de origem de planta, a lixiviação de minerais a partir do minério e o tratamento de esgoto e outros resíduos. As bactérias usadas nestes processos são modificadas por engenharia genética de modo que seu trabalho seja mais eficiente. Estão sendo desenvolvidas novas cepas de bactérias tecnologicamente úteis que vão degradar substâncias químicas tóxicas e poluentes, melhorar a recuperação de óleo, aumentar a captação de nitrogênio pelas plantas e inibir o crescimento de bactérias e fungos patogênicos.

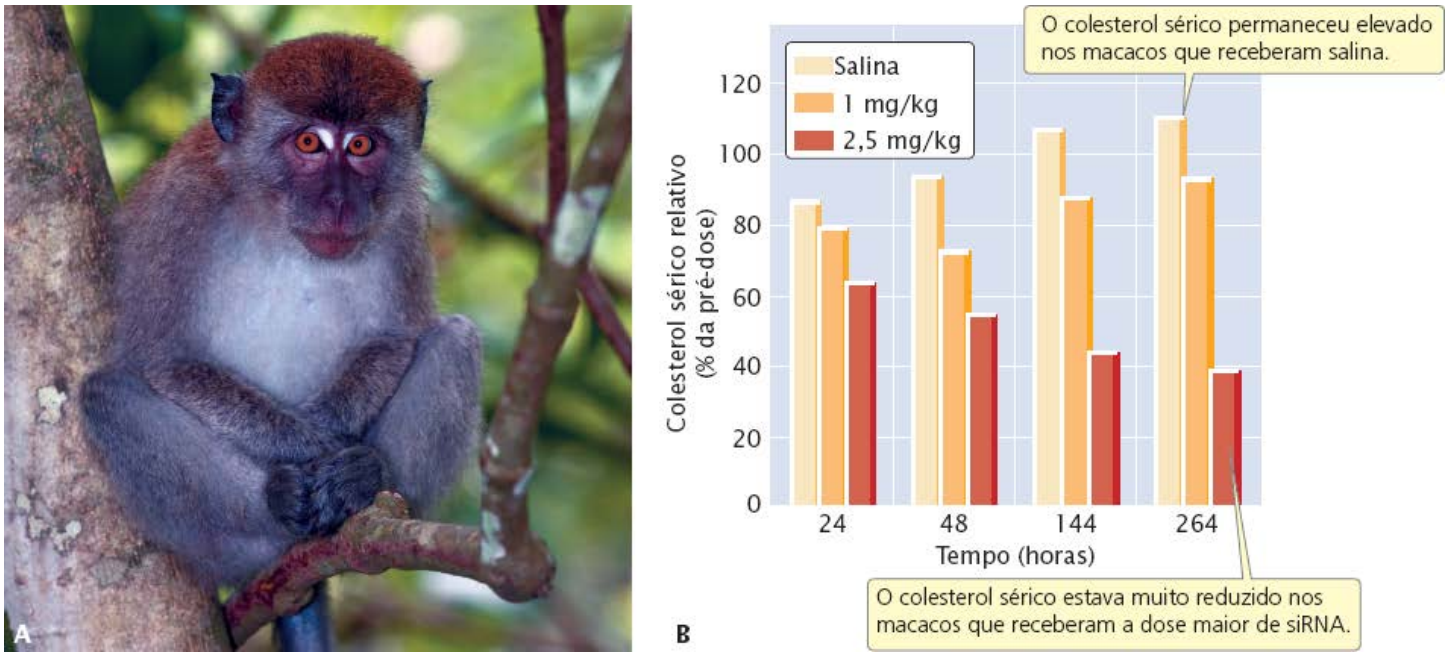


Figura 19.32 A transferência de siRNAs para a proteína ApoB nos macacos cinomolgo reduziu de forma significativa os níveis de colesterol no sangue. **A.** Macaco cinomolgo. **B.** Grupos diferentes de macacos receberam salina (controle), uma dose de 1 mg/kg ou uma dose de 2,5 mg/kg de siRNAs. (*Parte A*: Tony Camacho/Science Source. *Parte B*: De T. S. Zimmerman, Nature, 441:112, 2006, Figure 3b.)

Produtos agrícolas

A tecnologia do DNA recombinante tem um efeito importante na agricultura, onde é usada para criar plantações e animais domésticos com traços genéticos valiosos. Por muitos anos, os patologistas de plantas sabiam que as plantas infectadas com cepas atenuadas de vírus eram resistentes à infecção pelas cepas virulentas. Usando esse conhecimento, os geneticistas criaram resistência viral nas plantas ao transferir os genes para proteínas virais a células das plantas. Uma abóbora geneticamente modificada, chamada *Freedom* II, carreia genes para o vírus 2 do mosaico da melancia e o vírus do mosaico amarelo da abobrinha, que protege a abóbora contra infecções virais.

Outro objetivo era resistência a pragas usando a engenharia genética nas plantas para reduzir a dependência de pesticidas químicos. Como já foi discutido neste capítulo, um gene da bactéria *Bacillus thuringiensis* que produz uma toxina inseticida foi transferido para milho, tomate, batata, algodão e outras plantas. Atualmente, estas plantações Bt são cultivadas em todo mundo. Outros genes que conferem resistência a vírus e herbicidas foram introduzidos em várias plantações. Em 2011, 16,7 milhões de agricultores pelo mundo plantaram 160 milhões de hectares de plantações geneticamente modificadas. Nos EUA, 88% de todo milho, 94% de todo algodão e 93% da soja cultivada em 2012 eram geneticamente modificados.

As técnicas de DNA recombinante também se aplicam a animais domésticos. Por exemplo, o gene para o hormônio do crescimento foi isolado do gado e clonado na *E. coli*; estas bactérias produzem grandes quantidades do hormônio do crescimento bovino, que é administrado ao gado leiteiro para aumentar a produção de leite. Foram desenvolvidos animais transgênicos para carrear genes que codificam produtos farmacêuticos, algumas proteínas eucarióticas devem ser modificadas após a tradução e apenas outros eucariotos (mas não as bactérias) podem fazer as modificações. Por exemplo, um gene para o fator VIII de coagulação humano foi preso à região reguladora do gene do carneiro para β-lactoglobulina, uma proteína do leite. O gene fundido foi injetado em embriões de carneiro, criando carneiro transgênico que produzia no seu leite o fator de coagulação humano, que é usado para tratar hemofílicos. Foi criado salmão transgênico para carrear um gene do crescimento exógeno e o promotor; o peixe transgênico cresce o ano inteiro em vez de apenas durante os meses quentes, alcançando o tamanho de mercado mais rapidamente e com menos alimento do que salmão selvagem. Por meio da engenharia genética, os cientistas criaram galinhas transgênicas que expressam o RNA de pequeno tamanho molecular que bloqueia a infecção pelo vírus da gripe aviária.

A engenharia genética dos produtos de agricultura é controversa. Uma área de preocupação se dedica aos potenciais efeitos de liberar organismos novos produzidos pela engenharia genética no ambiente. Existem muitos exemplos nos quais organismos não nativos liberados em novos ambientes causaram alteração ecológica porque não têm predadores nem outros mecanismos de controle natural. Normalmente, a engenharia genética transfere apenas pequenas sequências de DNA, relativas a grandes diferenças genéticas que existem entre as espécies, mas até pequenas diferenças genéticas podem alterar traços genéticos ecologicamente importantes que podem afetar o ecossistema.

Outra área de preocupação é o efeito das culturas geneticamente modificadas sobre a biodiversidade. No maior teste de campo de plantas geneticamente modificadas já conduzido, os cientistas cultivaram beterraba, milho e colza que foram geneticamente modificados para resistir a herbicida junto com plantações tradicionais em 200 grupos de teste por todo o Reino Unido. Eles mediram a biodiversidade das plantas nativas e animais nos campos de cultivo. E descobriram que as plantas geneticamente modificadas conseguiam suprimir as pragas, entretanto, os grupos com beterrabas e colza geneticamente modificadas tinham bem menos insetos que se alimentavam das pragas. Por exemplo, os grupos com colza geneticamente modificada tinham 24% menos borboletas do que os grupos com plantações tradicionais.

Também existe a preocupação de que os organismos transgênicos possam hibridizar com os organismos nativos e transferir seus traços geneticamente modificados. Por exemplo, a resistência a herbicida modificada nas plantações pode ser transferida para as pragas, que seriam resistentes aos herbicidas que são usados para seu controle. Alguns estudos detectaram a hibridização entre culturas geneticamente modificadas e populações selvagens das plantas. Por exemplo, existe indício de que a colza transgênica (*Brassica napus*) tenha hibridizado com a *Brassica rapa* selvagem no Canadá. Outras preocupações se voltam para questões de segurança da saúde associadas à presença de produtos modificados em alimentos naturais; alguns críticos defendem a exigência de rotulagem de todos os alimentos geneticamente modificados que tenha DNA ou proteína transgênica. Esta rotulagem é exigida em países da União Europeia, mas não nos EUA.

Por outro lado, o uso de culturas e animais domésticos geneticamente modificados tem potenciais benefícios. As culturas geneticamente modificadas resistentes a pragas têm o potencial de reduzir o uso de substâncias químicas prejudiciais para o meio ambiente e os achados de pesquisas indicam que são usadas quantidades menores de pesticidas nos EUA como resultado da adoção de plantas transgênicas. Estudos conduzidos na China mostraram que quando as plantações com Bt são usadas, os fazendeiros pulverizam menos agrotóxicos químicos, permitindo que mais insetos predatórios sobrevivam e criando maior controle natural das pragas. As culturas transgênicas também aumentam os rendimentos, fornecendo mais alimentos por acre, o que reduz a quantidade de terra a ser usada pela agricultura. As plantas geneticamente modificadas oferecem o potencial para maior rendimento que pode ser necessário para alimentar a futura população mundial.

Conceitos

A tecnologia do DNA recombinante é usada para criar uma ampla gama de produtos químicos, incluindo produtos farmacêuticos, bactérias especializadas, culturas geneticamente modificadas e animais domésticos transgênicos.

Checagem dos conceitos 12

Quais são as questões envolvidas com as culturas geneticamente modificadas?

Exame genético

A identificação e clonagem de muitos genes responsáveis por doenças nos humanos permitiram o desenvolvimento de sondas para detectar mutações causadoras de doenças. Já está disponível o exame pré-natal para muitas doenças genéticas (ver *Capítulo 6*). Além disso, estão disponíveis testes genéticos pré-sintomáticos para adultos e crianças para vários distúrbios. Vários testes genéticos são oferecidos diretamente aos consumidores, sem exigir a participação de um profissional de saúde. Os testes genéticos de coleta domiciliares, em geral oferecidos pela internet, estão disponíveis para testar uma gama grande e crescente de condições genéticas, tudo desde distúrbios de um único gene como a fibrose cística a condições multifatoriais como obesidade, doença cardiovascular, desempenho atlético e predisposição para dependência a nicotina.

A crescente disponibilidade de testes genéticos levanta várias questões éticas e sociais. Por exemplo, é ético fazer um teste para doenças genéticas para as quais não existe cura ou tratamento? Outras questões éticas e jurídicas discutem a confidencialidade dos resultados do teste. Quem deve ter acesso aos resultados do exame genético? Os parentes que estão sob risco devem ser informados sobre os resultados do teste?

Outras questões estão relacionadas com a precisão dos testes genéticos. Para muitas doenças genéticas, os únicos testes preditivos disponíveis são os que identificam uma *predisposição* a mutação no DNA, mas muitas doenças genéticas podem ser causadas por dezenas ou centenas de diferentes mutações. Podem ser desenvolvidas sondas que detectam mutações comuns, mas elas não detectarão mutações raras e podem dar um resultado falso-negativo. Sem poder sequenciar o gene inteiro, o que seria caro e demorado, não existe uma forma de identificar todas as pessoas predispostas. Estas questões e preocupações são atualmente o foco de intenso debate por profissionais da área de ética, médicos, cientistas e pacientes.

Terapia gênica

A principal aplicação da terapia do DNA recombinante é a **terapia gênica**, a transferência direta de genes nos humanos para

tratar doenças. Atualmente, milhares de pacientes receberam esta terapia e muitos estudos clínicos estão em andamento. A terapia gênica é usada como tratamento experimental para doenças genéticas, câncer, cardiopatia e até algumas infecções como a AIDS. Estão em desenvolvimento vários métodos diferentes para transferir genes para as células humanas. Os vetores comumente usados incluem retrovírus, adenovírus e vírus adeno-associados geneticamente modificados (**Quadro 19.4**).

Apesar do número crescente de estudos clínicos para terapia gênica, existem problemas significativos para transferir genes exógenos para as células humanas, conseguir que sejam expressos e limitar as respostas imunológicas aos produtos dos genes e aos vetores usados para transferir os genes para as células. Também existem preocupações sobre a segurança. Em 1999, um paciente que participava de um estudo de terapia gênica teve uma reação imunológica fatal após receber uma injeção com o vetor viral que carreava um gene para tratar seu distúrbio metabólico. Além disso, cinco crianças submetidas à terapia gênica para imunodeficiência combinada desenvolveram leucemia que parecia estar diretamente relacionada com a inserção de gene de vetores retrovirais nos genes causadores do câncer. Apesar destes empecilhos, a terapia gênica continua progredindo. Foram publicados resultados claros que demonstram benefícios positivos a partir da terapia gênica para várias doenças diferentes (ver a introdução deste capítulo).

A terapia gênica conduzida até o momento se concentra apenas células não reprodutivas, somáticas. Corrigir o defeito genético nestas células (chamada de *terapia gênica somática*) pode fornecer benefícios positivos para os pacientes, mas não afetará os genes das futuras gerações. Esta terapia que altera as células reprodutivas, ou linhagem celular (chamada de *terapia gênica de linhagem celular*) é tecnicamente possível, mas levanta várias questões éticas porque tem a capacidade de alterar o conjunto de genes das futuras gerações.

Conceitos

A terapia gênica é a transferência direta dos genes nos humanos para tratar doenças. Ela foi implantada com sucesso em 1990 e agora é usada para tratar doenças genéticas, câncer e doenças infecciosas.

Checagem dos conceitos 13

Qual a diferença entre a terapia gênica somática e a terapia gênica de linhagem celular?

Quadro 19.4 Vetores usados na terapia gênica.

Vetor	Vantagens	Desvantagens
Retrovírus	Transferência eficiente	Transfere DNA apenas para as células em divisão, insere aleatoriamente o risco de produzir vírus do tipo selvagem
Adenovírus	Transfere para células que não estão em divisão	Provoca reação imunológica
Vírus associado ao adenovírus	Não provoca reação imunológica	Mantém pequena quantidade de DNA, difícil processar
Herpes-vírus	Pode ser inserido em células do sistema nervoso, não provoca reação imunológica	Difícil produzir em grandes quantidades
Lentivírus	Pode acomodar genes maiores	Questões de segurança
Lipossomos e outros vetores revestidos de lipídio	Sem replicação, não estimula reação imunológica	Baixa eficiência
Injeção direta	Sem replicação, direcionado para tecidos específicos	Baixa eficiência, não trabalha bem dentro de alguns tecidos
Tratamento de pressão	Seguro, porque os tecidos são	Mais eficiente com pequenas

	tratados fora do corpo e então transplantados para o paciente	moléculas de DNA
Gene *gun* (DNA revestido com pequenas partículas de ouro e disparado para o tecido)	Sem necessidade de vetor	Baixa eficiência

Fonte: De E. Marshall, Gene therapy´s growing pains, Science 269:1050-1055, 1995.

Resumo dos conceitos

- As endonucleases de restrição são enzimas que fazem cortes de fita dupla no DNA em sequências de base específicas
- Os fragmentos de DNA podem ser separados com o uso de eletroforese em gel e visualizados ao corar o gel com um corante específico para os ácidos nucleicos ou marcar os fragmentos com marcadores radioativos ou químicos
- Na clonagem de genes, um gene ou um fragmento de DNA é colocado em uma célula bacteriana, onde será multiplicado à medida que a célula se divide
- Os plasmídios, pequenos fragmentos circulares de DNA, são usados como vetores para garantir que um gene clonado seja estável e replicado dentro das células receptoras. Os vetores de expressão têm sequências necessárias para que o DNA exógeno seja transcrito e traduzido
- A reação em cadeia da polimerase é um método para amplificar enzimaticamente o DNA sem clonagem. Uma solução contendo DNA é aquecida, as duas fitas de DNA se separam e são rapidamente resfriadas, permitindo que os *primers* se prendam ao DNA molde. A solução é aquecida novamente e a DNA polimerase sintetiza novas fitas a partir dos *primers*. Cada vez que o ciclo é repetido, a quantidade de DNA dobra
- Os genes podem ser isolados ao criar uma biblioteca de DNA, um conjunto de colônias bacterianas ou placas virais cada uma com um diferente fragmento clonado de DNA. Uma biblioteca genômica tem o genoma inteiro de um organismo, uma biblioteca de cDNA tem fragmentos de DNA complementares a todos os mRNA diferentes em uma célula
- A hibridização *in situ* pode ser usada para determinar a localização cromossômica de um gene e a distribuição do mRNA produzido por um gene
- A clonagem posicional usa relações de ligação para determinar a localização dos genes sem qualquer conhecimento de seus produtos
- O método (didesoxi) de sequenciamento de DNA de Sanger usa substratos especiais para a síntese de DNA (trifosfatos de didesoxirribonucleosídios, ddNTPs) que terminam a síntese após serem incorporados ao DNA recém-sintetizado. Os métodos de sequenciamento de próxima geração e terceira geração sequenciam muitos fragmentos de DNA simultaneamente, fornecendo uma determinação mais rápida e mais barata de uma sequência de DNA
- As repetições curtas em *tandem* (STRs) e os microssatélites são usados para identificar as pessoas por meio de suas sequências de DNA (*fingerprinting* do DNA)
- A genética direta começa com um fenótipo e conduz análises para localizar os genes responsáveis. A genética reversa começa com uma sequência de DNA e conduz análises para determinar seu efeito fenotípico
- A mutagênese direcionada pelo sítio pode ser usada para produzir mutações em sítios específicos no DNA, permitindo que os genes sejam adaptados para um objetivo particular
- Os animais transgênicos, produzidos ao injetar DNA nos óvulos fertilizados, têm DNA estranho que é integrado a um cromossomo. Os camundongos *knockout* são camundongos nos quais um gene normal está desativado
- A interferência por RNA é usada para silenciar a expressão de genes específicos
- As técnicas de genética molecular estão sendo usadas para criar produtos de importância comercial, desenvolver testes de diagnóstico e tratar doenças
- Na terapia gênica, as doenças estão sendo tratadas ao alterar os genes das células humanas.

Termos importantes

Biblioteca de cDNA (DNA complementar)
Biblioteca de DNA
Biblioteca genômica
Biotecnologia
Caminhada cromossômica
Camundongo *knock-in*

Camundongo *knockout*
Clonagem de genes
Clonagem posicional
Cosmídio
Cromossomo artificial de levedura (YAC)
Cromossomo bacteriano artificial (BAC)
Eletroforese em gel
Endonuclease de restrição
Engenharia genética
Enzima de restrição
Extremidade coesiva
Fingerprinting do DNA
Genética direta
Genética reversa
Hibridização *in situ*
Ligador (*linker*)
Microssatélite
Mutagênese direcionada pelo sítio
Mutagênese direcionada por oligonucleotídio
Northern blotting
Nuclease modificada
PCR de transcrição reversa
PCR em tempo real
Plasmídio Ti
Polimorfismo de comprimento do fragmento de restrição (RFLP)
Reação em cadeia da polimerase (PCR)
Repetição curta em *tandem* (STR)
Sequenciamento de DNA
Sonda
Southern blotting
Taq polimerase
Tecnologia do DNA recombinante
Tecnologias de sequenciamento de próxima geração
Terapia gênica
Transgene
Trifosfato de didesoxirribonucleosídio (ddNTP)
Vetor de clonagem
Vetor de expressão
Western blotting

Respostas da Checagem dos conceitos

1. As enzimas de restrição existem naturalmente nas bactérias, que as usam para evitar a entrada de DNA viral.
2. c.
3. *Southern blotting* é usado para transferir DNA de um gel para um meio sólido. *Northern blotting* é usado para transferir RNA de um gel para um meio sólido, e o *Western blotting* é usado para transferir proteína de um gel para um meio sólido.
4. O gene e o plasmídio são cortados com a mesma enzima de restrição e misturados. A DNA ligase é usada para selar os cortes no esqueleto de açúcar-fosfato.
5. Uma enzima DNA polimerase estável no calor é importante para o sucesso da PCR porque a primeira etapa da reação requer que a solução seja aquecida entre 90° e 100°C para separar as duas fitas. A maioria das enzimas é desnaturada a esta temperatura. Com o uso de uma polimerase estável no calor, a enzima pode ser adicionada no início da reação e funcionará por múltiplos ciclos.
6. Com o uso do código genético e da sequência de aminoácidos da proteína, podem ser deduzidas possíveis sequências de nucleotídios que cobrem uma pequena região do gene. É usada uma mistura de todas as sequências de nucleotídios

possíveis que possam codificar a proteína, levando em consideração os códons sinônimos, para sondar a biblioteca. Para minimizar o número de sequências necessárias, é selecionada uma região da proteína com relativamente pouca degeneração nos seus códons.

7. O padrão de expressão do gene pode ser examinado e a região codificadora de cópias do gene dos indivíduos com o fenótipo mutante pode ser comparada com a região codificadora dos indivíduos do tipo selvagem.
8. d.
9. Ao usar PCR com *primers* que flanqueiam a região contendo as repetições em *tandem*.
10. b.
11. c.
12. As possíveis preocupações incluem: (a) dano ecológico causado pela introdução de novos organismos ao ambiente; (b) efeitos negativos dos organismos transgênicos na biodiversidade; (c) possível disseminação de transgenes a organismos nativos pela hibridização e (d) efeitos sobre a saúde ao comer alimentos geneticamente modificados.
13. A terapia gênica somática modifica os genes apenas no tecido somático e estas modificações não podem ser herdadas. A terapia gênica da linhagem celular altera os genes nas células da linhagem celular e estes serão herdados.

Problemas desenvolvidos

Problema 1

Uma molécula de DNA de fita dupla com 5 milhões de pares de base tem uma composição de base com 62% G + C. Quantas vezes, em média, é provável que os sítios de restrição a seguir estejam presentes nesta molécula de DNA?

a. *Bam*HI (a sequência de reconhecimento é GGATCC).
b. *Hin*dIII (a sequência de reconhecimento é AAGCTT).
c. *Hpa*II (a sequência de reconhecimento é CCGG).

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
O número de sítios de restrição presentes na molécula de DNA para cada enzima de restrição especificada.

Quais informações são fornecidas para solucionar o problema?

- O tamanho da molécula de DNA
- A composição das bases G + C na molécula de DNA
- As sequências de reconhecimento para cada enzima de restrição.

Para a solução deste problema, revise:
Corte e união dos fragmentos de DNA, na Seção 19.2.

Etapas para a solução

As porcentagens de G e C são iguais no DNA de fita dupla, então se G + C = 62%, logo %G = %C = 62%/2 = 31%.

> Dica: Se você conhecer a porcentagem de qualquer base no DNA, você pode determinar as porcentagens de todas as outras bases, pois G = C e A = T.

A porcentagem de A + T = (100% – G – C) = 38% e %A = %T = 38%/2 = 19%. Para determinar a probabilidade de encontrar uma sequência de base específica, usamos a regra de multiplicação, multiplicando a probabilidade de encontrar cada base em um sítio específico.

> Lembrete: A regra de multiplicação afirma que a probabilidade de dois ou mais eventos independentes é calculada ao multiplicar suas probabilidades independentes.

a. A probabilidade de encontrar a sequência GGATCC = 0,31 × 0,31 × 0,19 × 0,19 × 0,31 × 0,31 = 0,0003333. Para determinar o número médio de sequências de reconhecimento em um fragmento de DNA de 5 milhões de pares de base, multiplicamos 5.000.000 pb × 0,00033 = 1.666,5 sequências de reconhecimento.
b. O número de sequências de reconhecimento AAGCTT é = 0,19 × 0,19 × 0,31 × 0,31 × 0,19 × 0,19 ×

5.000.000 = 626 sequências de reconhecimento.

c. O número de sequências de reconhecimento CCGG é = 0,31 × 0,31 × 0,31 × 0,31 × 5.000.000 = 46.176 sequências de reconhecimento.

Problema 2

Você recebe o seguinte fragmento de DNA para sequenciar: 5′-GCTTAGCATC-3′. Primeiro você clona o fragmento nas células bacterianas para produzir DNA suficiente para sequenciar. Depois, isola o DNA das células bacterianas e usa o método de sequenciamento didesoxi. E, por fim, separa os produtos das reações de polimerização por eletroforese a gel. Desenhe as bandas que devem aparecer no gel a partir das quatro reações de sequenciamento.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
As posições das bandas no gel de sequenciamento.

Quais informações são fornecidas para solucionar o problema?
A sequência de bases do fragmento de DNA a ser sequenciado.

Para a solução deste problema, revise:
Sequenciamento de DNA, na Seção 19.4.

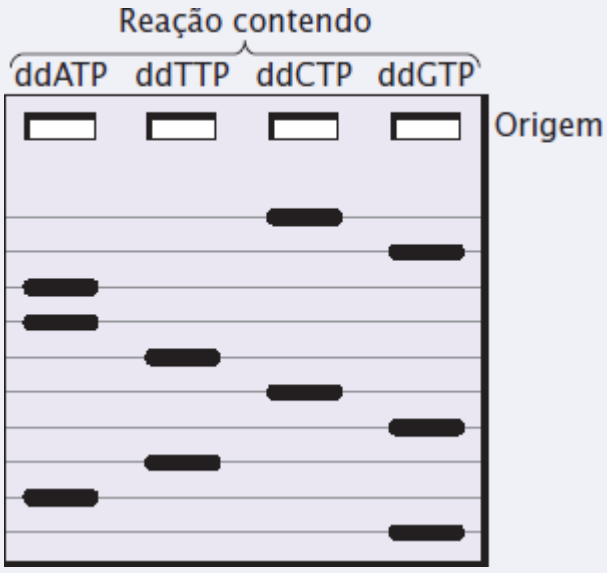
Etapas para a solução

A primeira etapa é escrever a sequência do fragmento recém-sintetizado, que será complementar e antiparalelo ao fragmento original. A sequência original é 5′-GCTTAGCATC-3′, então a sequência recém-sintetizada será:

Dica: Fragmentos pequenos, os mais próximos da extremidade 5′ da fita recém-sintetizada, vão migrar mais rápido e aparecerão próximo da porção inferior do gel.

Sequência original (molde): 5′-GCTTAGCATC-3′
Fita recém-sintetizada: 3′-CGAATCGTAG-5′
Assim, a sequência da fita recém-sintetizada, escrita de 5′ × 3′ é: 5′-GATGCTAAGC-3′. As bandas que representam esta sequência vão aparecer no gel, com as bandas representando os nucleotídios próximos da extremidade 5′ da molécula na parte inferior do gel.

Lembrete: No sequenciamento didesoxi, uma nova fita de DNA é sintetizada e esta fita é a que será sequenciada. Assim, as bandas que aparecem no gel representam o complemento da sequência original.



Questões de compreensão

Seção 19.1

1. Liste alguns dos efeitos e aplicações práticas das análises genéticas moleculares.

Seção 19.2

2. Qual característica é comumente observada nas sequências reconhecidas pelas enzimas de restrição do tipo II?
3. Qual é o papel normal as enzimas de restrição nas bactérias? Como as bactérias protegem seu próprio DNA da ação das enzimas de restrição?
4. Explique como a eletroforese em gel é usada para separar os fragmentos de DNA de diferentes comprimentos.
5. Como os fragmentos de DNA podem ser visualizados após serem separados por eletroforese em gel?
6. Qual é o objetivo do *Southern blotting*? Como é seu processo?
7. Indique três características importantes dos vetores de clonagem.
8. Descreva resumidamente dois métodos diferentes para inserir DNA exógeno nos plasmídios, indicando as vantagens e desvantagens de cada método.
9. Explique resumidamente como um gene para resistência a antibiótico e o gene *lacZ* podem ser usados para determinar quais células têm um plasmídio específico?
10. Explique resumidamente como a reação em cadeia da polimerase é usada para ampliar uma sequência específica de DNA. Quais são algumas das limitações da PCR?
11. O que é PCR em tempo real?

Seção 19.3

12. Qual a diferença entre biblioteca genômica e biblioteca de cDNA?
13. Como sondas são usadas para rastrear as bibliotecas de DNA? Explique como uma sonda sintética pode ser preparada quando se conhece a proteína de um gene.
14. Explique resumidamente a hibridização *in situ*, dando algumas aplicações desta técnica.
15. Explique resumidamente como um gene pode ser isolado por clonagem posicional.
16. Explique como a caminhada cromossômica pode ser usada para encontrar um gene.

Seção 19.4

17. Qual o objetivo do trifosfato de didesoxirribonucleosídio na reação de sequenciamento didesoxi?
18. O que é *fingerprinting* do DNA? Quais tipos de sequências são examinadas no *fingerprinting* do DNA?

Seção 19.5

19. Qual a diferença da abordagem da genética reversa para a abordagem da genética direta?
20. Explique resumidamente como é feita a mutagênese direcionada pelo sítio.
21. O que são camundongos *knockout*, como são produzidos e qual sua utilidade?
22. Como a interferência por RNA é usada na análise da função dos genes?

Seção 19.6

23. O que é terapia gênica?

Questões e problemas aplicados

Introdução

24. A introdução deste capítulo explica como a terapia gênica é usada para tratar a amaurose congênita de Leber (ACL, um tipo de cegueira grave). Quais características da cegueira que afetam a retina (como a ACL) que a tornam um candidato atraente para a terapia gênica?

Seção 19.2

*25. Suponha que um geneticista descobriu uma nova enzima de restrição na bactéria *Aeromonas ranidae*. Esta enzima de restrição é a primeira a ser isolada desta espécie de bactéria. Usando a convenção padrão para abreviar as enzimas de restrição, indique um nome para esta enzima de restrição (consulte o rodapé do Quadro 19.1).
26. Com que frequência, em média, você esperaria que uma endonuclease de restrição do tipo II cortaria uma molécula de DNA se a sequência de reconhecimento para enzima tiver 5 pb? (Considere que a probabilidade de encontrar os quatro

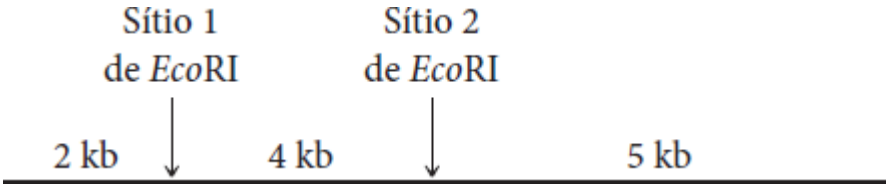
tipos de base no DNA é a mesma e que as bases em uma sequência de reconhecimento são independentes.) Com que frequência a endonuclease cortaria o DNA se a sequência de reconhecimento tivesse 8 pb?

*27. Um microbiologista descobre uma nova endonuclease de restrição do tipo II. Quando o DNA é digerido por esta enzima, são produzidos fragmentos com comprimento médio de 1.048.500 pb. Qual é o mais provável número de pares de base na sequência de reconhecimento desta enzima?

28. Os sítios de restrição para uma enzima com 4 pb no seu sítio de restrição estarão mais próximos, afastados ou na mesma distância, em média, comparados com os sítios de uma enzima com 6 pb no seu sítio de restrição? Explique sua justificativa.

*29. Cerca de 60% dos pares de base em uma molécula de DNA humano é AT. Se o genoma humano tem 3,2 bilhões de pares de bases de DNA, quantas vezes os sítios de restrição a seguir serão encontrados?
*Bam*HI (o sítio de restrição é 5′-GGATCC-3′)
*Eco*RI (o sítio de restrição é 5′-GAATTC-3′)
*Hae*III (o sítio de restrição é 5′-GGCC-3′)

*30. O mapeamento de restrição de um fragmento linear de DNA revela os seguintes sítios de restrição *Eco*RI.



a. Este segmento de DNA é cortado pela *Eco*RI, os fragmentos resultantes são separados por eletroforese em gel e o gel é corado com brometo de etídio. Desenhe uma imagem das bandas que aparecerão no gel.
b. Se ocorre uma mutação que altera o sítio 1 de *Eco*RI neste fragmento de DNA, qual será a diferença do padrão de bandeamento no gel a partir do que você desenhou na parte *a*?
c. Se ocorrem mutações que alteram os sítios 1 e 2 de *Eco*RI neste fragmento de DNA, qual será a diferença do padrão de bandeamento no gel a partir do que você desenhou na parte *a*?
d. Se forem inseridos 1.000 pb de DNA entre os dois sítios de restrição, qual será a diferença do padrão de bandeamento no gel a partir do que você desenhou na parte *a*?
e. Se forem eliminados 500 pb de DNA entre os dois sítios de restrição, qual será a diferença do padrão de bandeamento no gel a partir do que você desenhou na parte *a*?

*31. Quais vetores (plasmídio, fago 1, cosmídio, cromossomo bacteriano artificial) podem ser usados para clonar um fragmento contínuo de DNA com os seguintes comprimentos?
a. 4 kb.
b. 20 kb.
c. 35 kb.
d. 100 kb.

32. Um geneticista usa um plasmídio para clonagem que tem o gene *lacZ* e um gene que confere resistência à penicilina. Ele insere um fragmento de DNA exógeno em um sítio de restrição localizado entre o gene *lacZ* e usa o plasmídio para transformar as bactérias. Explique como o geneticista pode identificar as bactérias que têm uma cópia de um plasmídio com o DNA exógeno.

33. Na Figura 19.10, qual o objetivo de o gene neo^+ estar ligado ao gene *Bt*?

Seção 19.3

34. Suponha que você terminou a faculdade e começou a trabalhar em uma empresa de biotecnologia. Seu primeiro trabalho é clonar o gene do porco para o hormônio prolactina. Considere que o gene do porco para prolactina ainda não foi isolado, sequenciado ou mapeado; entretanto, o gene do camundongo para prolactina foi clonado e a sequência de aminoácidos desta prolactina é conhecida. Explique resumidamente duas estratégias diferentes que você pode usar para encontrar e clonar o gene do porco para prolactina.

*35. Um biólogo molecular quer isolar um gene de um escorpião que codifica uma toxina mortal encontrada no seu ferrão, com o objetivo de transferir seu gene para as bactérias e produzir a toxina para uso como pesticida comercial. O isolamento do gene exige uma biblioteca de DNA. O biólogo molecular deve criar uma biblioteca genômica ou uma biblioteca de cDNA? Explique sua justificativa.

*36. Uma proteína tem a seguinte sequência de aminoácidos:

Met-Tyr-Asn-Val-Arg-Val-Tyr-Lys-Ala-Lys-Trp-Leu-Ile-His-Thr-Pro

Você quer fazer um conjunto de sondas para rastrear uma biblioteca de cDNA à procura da sequência que codifica esta proteína. Suas sondas devem ter pelo menos 18 nucleotídios de comprimento.

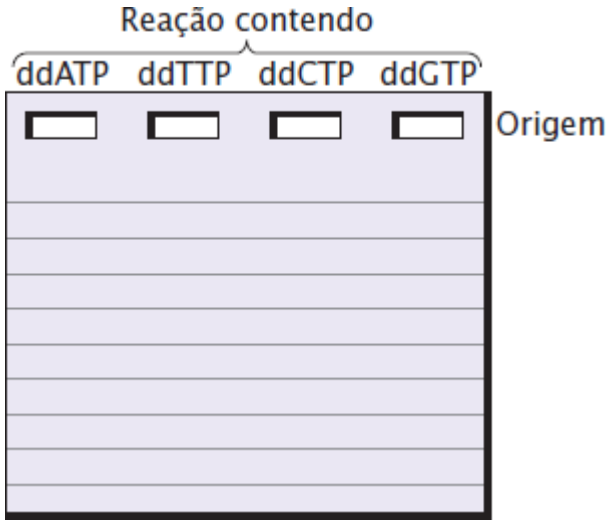
a. Quais aminoácidos na proteína devem ser usados para construir as sondas de modo que resulte o mínimo em degeneração? (Consulte o código genético na Figura 15.8.)
b. Quantas sondas diferentes precisam ser sintetizadas para ter certeza de que você encontrará a sequência correta de cDNA que especifica a proteína?

Seção 19.4

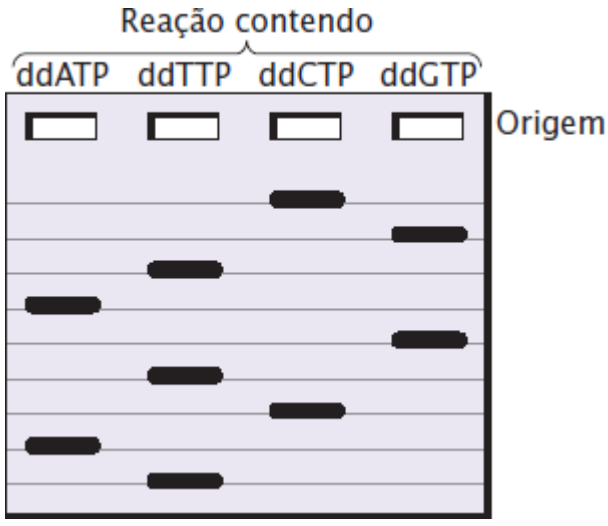
37. Na Figura 19.19, Bob e Joe são homozigotos para diferentes padrões de fragmento de restrição. Quantas bandas você esperaria ver no gel se uma pessoa for heterozigota para os padrões A e B? Explique sua justificativa.
38. Suponha que você vai sequenciar o seguinte fragmento de DNA:

5′-TCCCGGGAAA-sítio do *primer*-3′

Primeiro você usa PCR para amplificar o fragmento, então existe DNA suficiente para sequenciamento. Separe os produtos das reações de polimerização por eletroforese em gel. Desenhe as bandas que devem aparecer no gel a partir das quatro reações de sequenciamento.



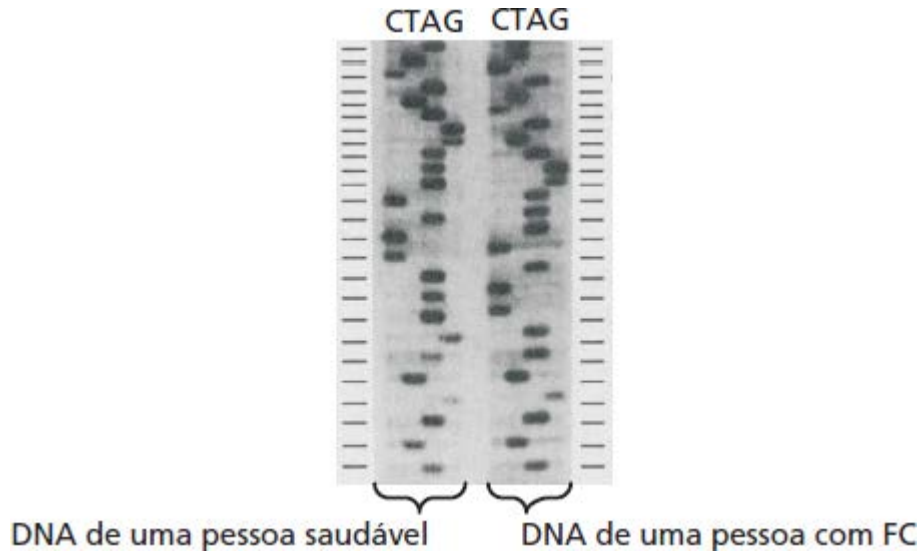
*39. Suponha que você recebeu um pequeno fragmento de DNA para sequenciar. Primeiro, você amplia o fragmento com PCR e ajusta uma série de quatro reações didesoxi. Depois, separa os produtos das reações por eletroforese a gel e obtém o seguinte padrão de bandeamento:



Escreva a sequência de bases do fragmento original que você recebeu.

Sequência original: 5′-_____-3′

40. A imagem a seguir é um gel de sequenciamento do estudo original que sequenciou pela primeira vez o gene da fibrose cística (J. R. Riordan *et al*. 1989. *Science* 245:1066-1073). A partir da imagem, determine a sequência da cópia normal do gene e a sequência da cópia mutante do gene. Identifique a localização da mutação responsável pela fibrose cística (FC). Dica: A mutação FC é uma deleção de 3 pb.

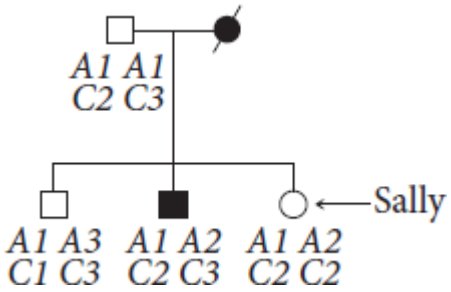


(De J. R. Riordan et al. Science 245:1066-1073. Reimpressa com permissão de AAAS.)

*41. Um distúrbio hipotético chamado de síndrome G é dominante autossômico e se caracteriza por defeitos visuais, esqueléticos e cardiovasculares. A condição surge na meia-idade. Como seus sinais/sintomas são variados, é difícil diagnosticá-la. O diagnóstico precoce é importante, entretanto, porque os defeitos cardiovasculares podem ser tratados se o distúrbio for identificado precocemente. Sabe-se que o gene para a síndrome G reside no cromossomo 7, e está intimamente ligado a dois RFLPs no mesmo cromossomo, um no *locus* A e outro no *locus* C. Os *loci G*, *A* e *C* estão muito próximos, e ocorre pouco *crossing over* entre eles. Os seguintes alelos de RFLP são encontrados nos *loci A* e *C*:

Locus A: *A1, A2, A3*
Locus C: *C1, C2, C3*

Sally, apresentada no heredograma a seguir, está preocupada se ela poderia ter a síndrome G. Ela perdeu sua mãe por causa da síndrome e tem um irmão já acometido pelo distúrbio. Seu outro irmão está na meia-idade e não tem a doença; então considere que ele não é carreador dos genes desta síndrome. Um geneticista faz o genótipo de Sally e de seus familiares próximos para os *loci A* e *C* e obtém os genótipos apresentados na linhagem.



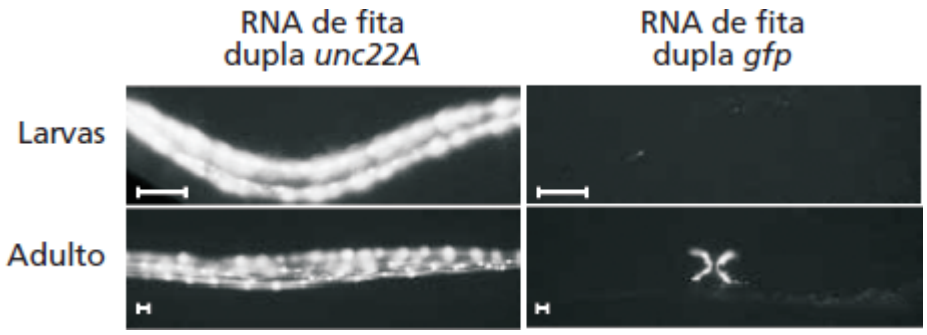
a. Considere que não ocorre *crossing over* entre os *loci* A, C e G. Sally tem o gene para a síndrome G? Por que sim ou por que não?
b. Desenhe a disposição dos alelos *A*, *C* e *G* nos cromossomos para todos os membros da família.

Seção 19.5

42. Você descobriu um gene nos camundongos que é semelhante a um gene na levedura. Como você poderia determinar se este gene é essencial para o desenvolvimento nos camundongos?

*43. Andrew Fire, Craig Mello e seus colaboradores estão entre os primeiros pesquisadores a examinar os defeitos do RNA de fita dupla na expressão gênica (A. Fire *et al*. 1998. *Nature* 391:806-811). Em um experimento, eles usaram uma cepa

transgênica de *C. elegans* no qual foi introduzido um gene (*gfp*) para um pigmento fluorescente verde. Eles injetaram nos vermes RNA de fita dupla complementar às sequências codificadoras do gene *gfp* e outros vermes receberam injeção de RNA de fita dupla complementar à região codificadora de um gene diferente (*unc22A*) que codifica uma proteína do músculo. As fotografias a seguir mostram as larvas e os descendentes adultos dos vermes que receberam a injeção. O pigmento fluorescente verde aparece como pontos brilhantes nas fotografias.



(Reimpressa com permissão de Macmillan Publishers Ltd: NATURE 391:806-811 [19 February 1998], copyright 1998.)

a. Explique estes resultados.
b. Fire e Mello conduziram outro experimento no qual eles injetaram RNA de fita dupla complementar aos íntrons e sequências de promotor do gene *gfp*. Quais resultados você esperaria com este experimento? Explique sua resposta.

Questões desafiadoras

Seção 19.5

44. Suponha que você foi contratado por uma empresa de biotecnologia para produzir uma cepa gigante de moscas-da-fruta usando a tecnologia do DNA recombinante de modo que os estudantes de genética não sejam obrigados a forçar a visão ao olhar para os minúsculos insetos. Você vai até a biblioteca e aprende que o crescimento nas moscas-da-fruta é normalmente inibido por um hormônio chamado de substância P pequena (SSP). Você decide que pode produzir moscas-da-fruta gigantes se conseguir desligar de alguma maneira a produção de SSP. A substância P pequena é sintetizada por um composto chamado XSP em uma reação de uma etapa catalisada pela enzima runtase:

$$\text{XSP} \xrightarrow[\text{Runtase}]{} \text{SSP}$$

Um pesquisador isolou o cDNA para runtase e o sequenciou, mas a localização do gene da runtase no genoma da *Drosophila* não é conhecida. Ao tentar planejar uma estratégia para desligar a produção de SSP e produzir moscas gigantes ao usar as técnicas padrão de DNA recombinante, você descobre que será muito difícil eliminar, inativar ou induzir mutação nesta sequência de DNA na *Drosophila*. Portanto, você é obrigado a restringir sua interferência genética ao aumento do gene (adicionar novos genes às células). Descreva os métodos que você usará para desligar o SSP e produzir moscas gigantes ao usar a tecnologia do DNA recombinante.

Seção 19.6

45. Parte da controvérsia sobre os alimentos geneticamente modificados se concentra em se deve ser necessária uma rotulagem especial em todos os produtos preparados a partir de plantações geneticamente modificadas. Algumas pessoas defendem uma rotulagem que identifica o produto como sendo preparado a partir de plantas geneticamente modificadas. Outras argumentam que a rotulagem deve ser necessária apenas para identificar os ingredientes, não o processo pelo qual foram produzidos. Escolha um lado nessa e justifique sua opinião.

20 Genômica e Proteômica

Como decifrar a dança das abelhas | O genoma da abelha

A abelha, *Apis mellifera*, é um dos animais mais incríveis do mundo. As abelhas ão extremamente sociais, vivem em colônias complexas nas quais cooperam e assumem responsabilidade por tarefas especializadas que beneficiam a colmeia toda. Apesar de um cérebro minúsculo que tem apenas um milhão de neurônios (o cérebro humano tem 100 bilhões de neurônios), as abelhas têm comportamento intricado e se comunicam efetivamente com milhares de outras abelhas que compõem a colmeia. Cada abelha reconhece o odor de sua própria colmeia, assim como odores que sinalizam alarme, indicam tarefas a serem realizadas e distinguem a casta, a idade e o sexo de outras abelhas. As abelhas operárias têm uma memória extraordinária. Ao encontrar um grupo de flores que fornece um rico suprimento de néctar, uma abelha operária retorna para a colmeia e "dança", fazendo um oito no lado do favo de mel. A dança fornece informações para outras abelhas sobre a localização da fonte de alimento, incluindo a direção na referência ao ângulo do sol e a distância da colmeia para a fonte. Uma abelha consegue lembrar-se de cinco diferentes localizações de flores, incluindo a direção e a distância de cada local em relação à colmeia, pontos de referência ao longo do caminho e a hora do dia quando a flor produz mais néctar.

As abelhas não são apenas objetos de curiosidade biológica, elas têm um papel crítico na polinização de muitas plantações importantes, além da produção de mel e cera. Por causa da sua importância para a agricultura e a ciência, a abelha foi selecionada para sequenciamento pelo National Genome Research Institute do U.S. National Institutes of Health. A sequência completa do genoma foi publicada em 2006 e forneceu numerosas informações sobre o comportamento, a função neurológica, a ecologia e a evolução das abelhas.

Seu genoma é composto por 10.157 genes que incluem cerca de 236 milhões de pares de bases de DNA. Com base na informação genômica, foram identificados 36 genes que codificam neuropeptídios, proteínas do cérebro, que afetam o comportamento e a memória. Mais de 3.000 genes estão ativos no cérebro de uma abelha, e os pesquisadores localizaram regiões reguladoras que controlam a expressão de muitos destes genes, incluindo alguns que são importantes no desenvolvimento do comportamento de procura por alimento. Comparada com outros insetos, as abelhas têm muito mais genes codificadores de receptores olfatórios, explicando seu sofisticado sistema químico de comunicação.

Há muito tempo que a abelha está associada aos humanos, mas sua origem evolutiva precisa era desconhecida até recentemente. Com as informações obtidas da sequência genômica, os geneticistas usaram os polimorfismos de únicos nucleotídios (SNPs, sítios que variam na base existente em um único nucleotídio) para estudar as relações evolutivas entre abelhas de diferentes partes do mundo. Este estudo revelou que *A. mellifera* se originou na África e não na Ásia, como se acreditava. As populações africanas se expandiram para Europa e Ásia em duas ondas, uma onda populacional para Europa Ocidental e outra se expandindo para Ásia e Europa Oriental.

As abelhas europeias foram introduzidas na América do Norte pelos colonizadores europeus, mas algumas foram substituídas recentemente por descendentes das "abelhas assassinas" africanas, conhecidas por seu comportamento muito agressivo. Estas abelhas foram introduzidas no Brasil em 1956 e se espalharam para o norte e chegaram aos EUA. A análise dos SNPs das abelhas coletadas nos EUA antes de 1990 mostrou que elas tinham apenas os genes europeus, mas de 1993 a 2001, as abelhas do sul do Texas mostravam uma transição para os genes predominantemente de origem africana.

O sequenciamento do genoma das abelhas ilustra as diversas formas que a informação genômica está sendo usada atualmente. A **genômica** é o campo da genética que tenta compreender o conteúdo, a organização, a função e a evolução da informação genética contida em genomas inteiros. O campo da genômica é a pesquisa de ponta da biologia moderna, as informações oriundas deste campo fizeram contribuições significativas a saúde humana, agricultura e várias outras áreas. Ela forneceu sequências gênicas necessárias para produzir proteínas importantes para medicina por meio da tecnologia do DNA recombinante e as comparações de sequências genômicas provenientes de diferentes organismos estão levando a uma melhor compreensão da evolução e da história da vida.

Vamos começar este capítulo examinando os mapas genéticos e físicos e os métodos para sequenciar genomas inteiros. A seguir, vamos explorar a genômica funcional, como os genes são identificados nas sequências genômicas e como suas funções são definidas. A sequência de um genoma, por si só, tem pouco uso e agora que o sequenciamento de genomas se tornou rotineiro, boa parte da genômica está voltada para decifrar a função das sequências obtidas. Na parte final do capítulo vamos analisar a proteômica, o estudo do conjunto completo de proteínas encontrado em uma célula.

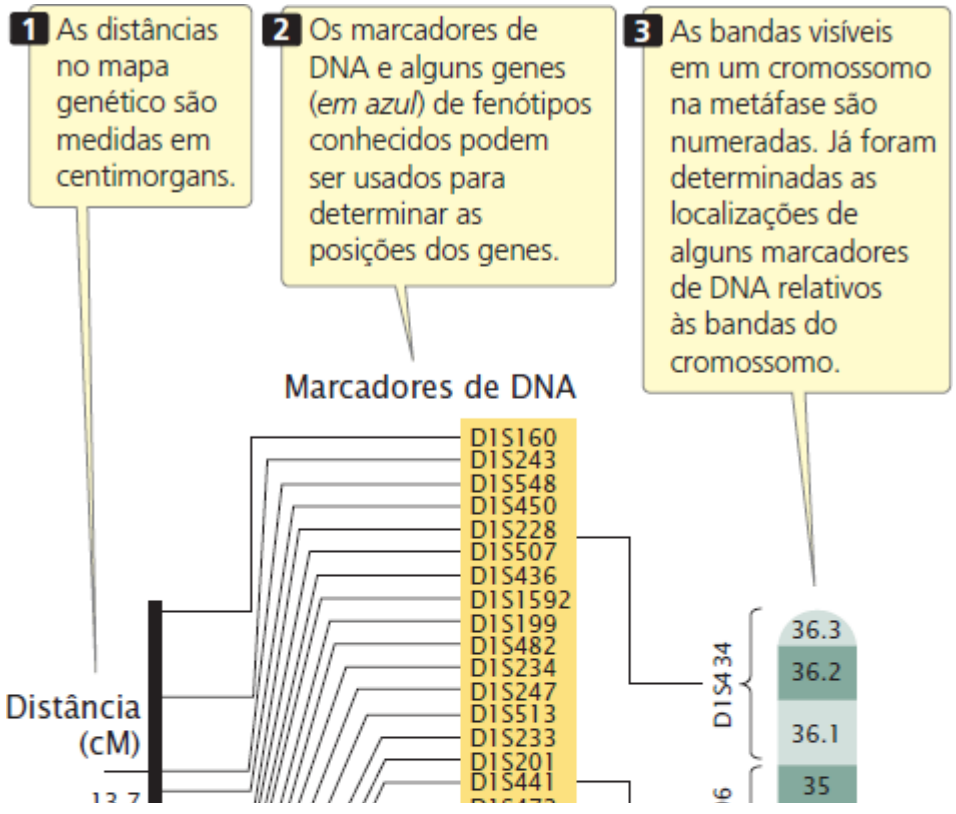
20.1 A genômica estrutural determina as sequências de DNA de genomas inteiros

A **genômica estrutural** estuda a organização e a sequência de informação genética contida em um genoma. Muitas vezes, uma etapa anterior na caracterização de um genoma consiste na preparação de mapas genéticos e físicos dos seus cromossomos. Estes mapas fornecem informações sobre as localizações relativas dos genes, marcadores moleculares e segmentos cromossômicos, que são essenciais para o posicionamento dos segmentos cromossômicos e alinhar trechos de DNA sequenciado em uma sequência de genoma inteiro.

Mapas genéticos

Todo mundo já usou um mapa uma vez na vida. Os mapas são indispensáveis para encontrar a casa de um amigo, um caminho em uma cidade estranha no seu estado ou a localização de um país. Cada um destes exemplos requer um mapa com uma diferente escala. Para encontrar a casa de um amigo, você usaria o mapa de ruas da cidade, para encontrar um caminho em uma cidade estranha, você pegaria um mapa das principais vias, e você precisaria de um atlas para encontrar um país como o Cazaquistão. Da mesma forma, navegar em um genoma requer mapas de diferentes tipos e escalas.

Construção dos mapas genéticos. Os **mapas genéticos** (também conhecidos como mapas de ligação) fornecem uma estimativa das localizações dos genes em relação a localizações de outros genes conhecidos (**Figura 20.1**). Estes mapas baseiam-se na função genética da recombinação (por isso, o nome mapa genético). Os princípios básicos da construção dos mapas genéticos foram discutidos em detalhes no Capítulo 7. Em resumo, organismos com genótipo conhecido são cruzados e a frequência de recombinação entre os *loci* é determinada pelo exame dos descendentes. Se a frequência de recombinação entre os *loci* for 50%, então os *loci* estão em diferentes cromossomos ou distantes no mesmo cromossomo. Se a frequência de recombinação for menor que 50%, os *loci* estão próximos no mesmo cromossomo (eles pertencem ao mesmo grupo de ligação). Para genes ligados, a taxa de recombinação é proporcional à distância física entre os *loci*. As distâncias nos mapas genéticos são medidas em porcentagem de recombinação (centimorgans, cM) ou unidades de mapa (m.u.). Os dados de cruzamentos múltiplos de dois ou três pontos podem ser integrados em mapas de ligação para cromossomos inteiros.



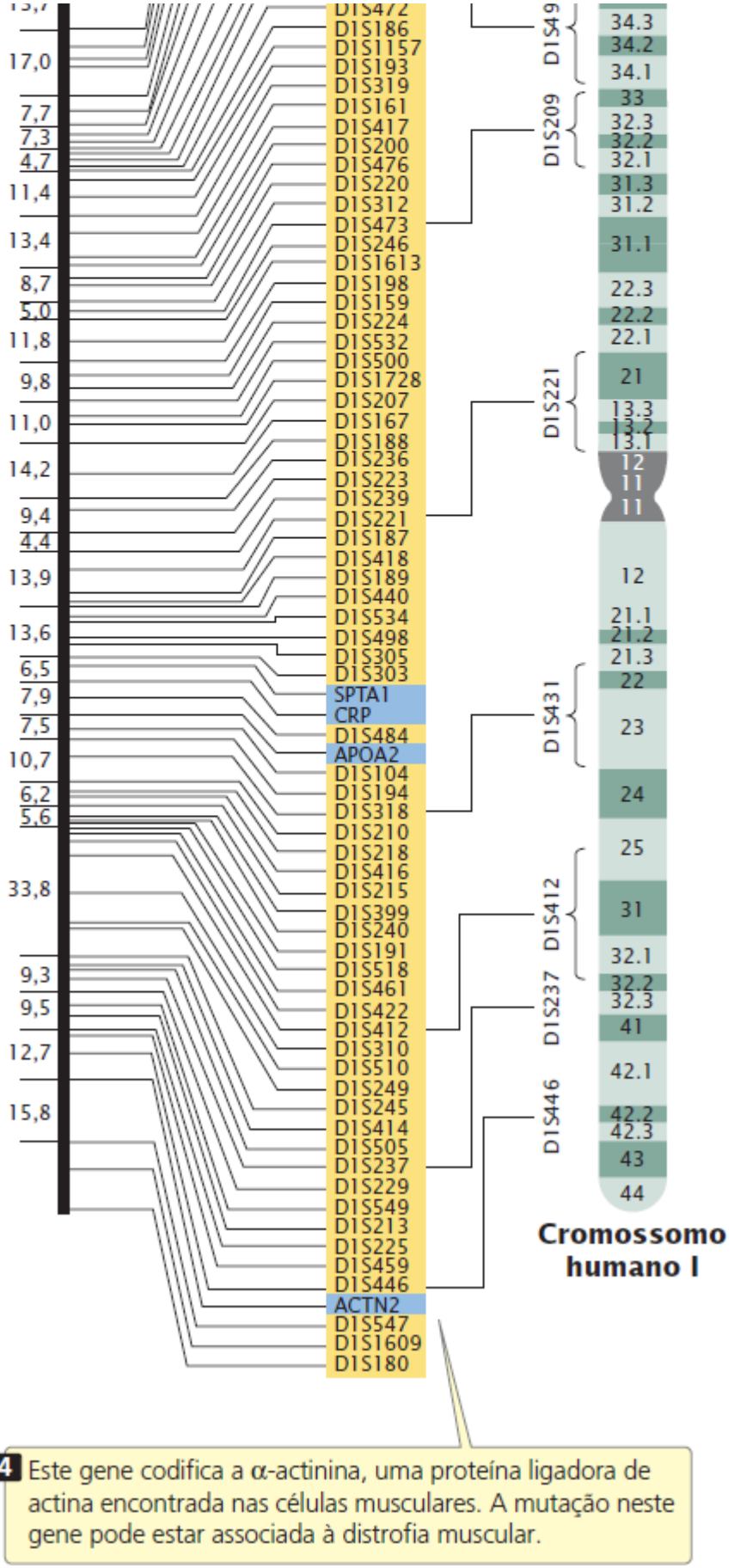


Figura 20.1 Os mapas genéticos baseiam-se em taxas de recombinação. Aqui está representado um mapa genético do cromossomo 1 humano.

Por muitos anos, os genes podiam ser detectados apenas pela observação de sua influência em um traço (o fenótipo) e a construção dos mapas genéticos era limitada pela disponibilidade de traços de *locus* único que poderiam ser examinados quanto a evidência de recombinação. Essa limitação acabou sendo superada pelo desenvolvimento de técnicas moleculares, como a análise dos polimorfismos de comprimento do fragmento de restrição, a reação em cadeia da polimerase e o sequenciamento do

DNA (ver *Capítulo 19*), que conseguem fornecer marcadores moleculares que podem ser usados para construir e refinar mapas genéticos.

Limitações dos mapas genéticos. Os mapas genéticos têm muitas limitações; a primeira delas é a resolução ou o detalhe. O genoma humano inclui 3,2 bilhões de pares de bases de DNA e tem uma distância genética total de 4.000 cM, uma média de 800.000 pb/cM. Mesmo que exista um marcador a cada centimorgan (o que não é uma expectativa realista), a resolução em relação à estrutura física do DNA ainda seria muito baixa. Em outras palavras, os detalhes do mapa seriam muito limitados. Um segundo problema dos mapas genéticos é que nem sempre eles correspondem exatamente às distâncias físicas entre os genes. Os mapas genéticos baseiam-se nas taxas de *crossing over*, que variam de uma parte de um cromossomo para outro, portanto, as distâncias em um mapa genético são apenas aproximações das distâncias físicas reais ao longo de um cromossomo. A **Figura 20.2** compara o mapa genético do cromossomo III de levedura com um mapa físico determinado por sequenciamento de DNA. Existem algumas discrepâncias entre as distâncias e até entre as posições de alguns genes. Apesar destas limitações, os mapas genéticos são críticos para o desenvolvimento dos mapas físicos e o sequenciamento dos genomas inteiros.

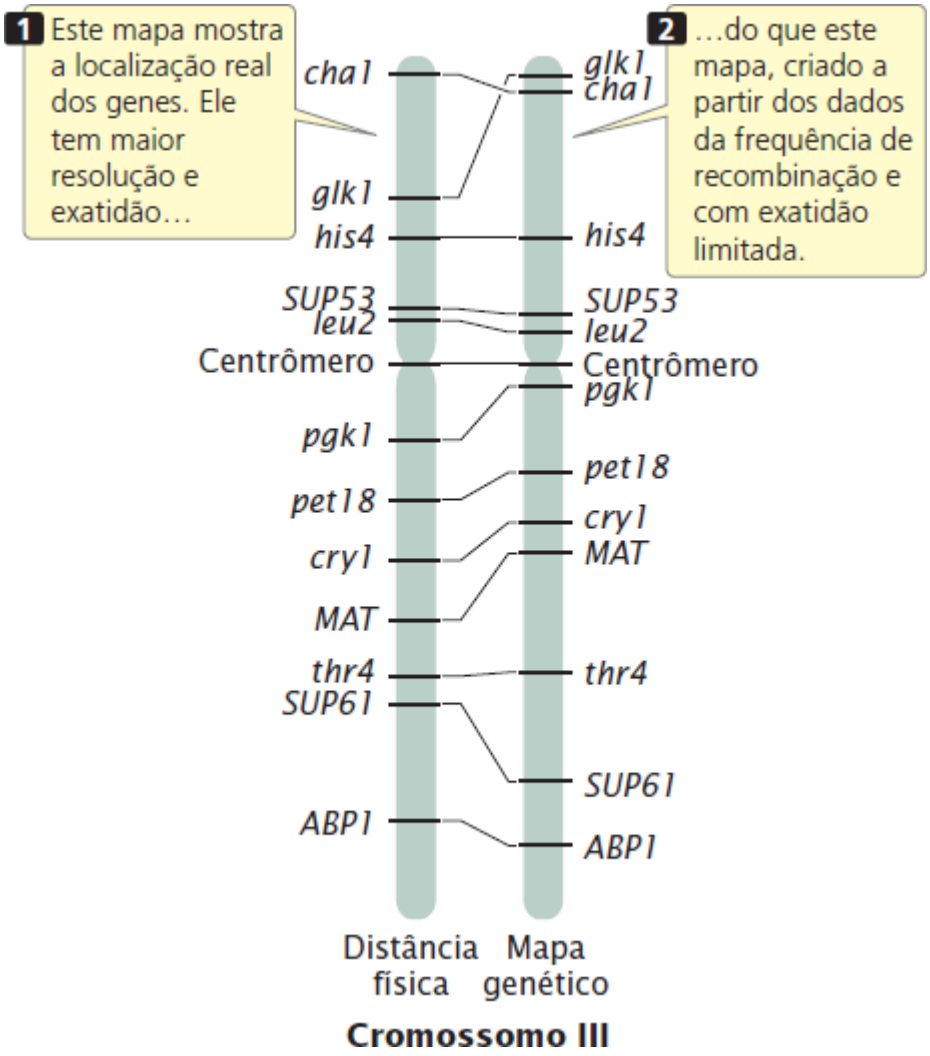


Figura 20.2 Os mapas genéticos e físicos podem apresentar diferenças nas distâncias relativas e até na posição dos genes em um cromossomo. Os mapas genéticos e físicos do cromossomo III da levedura revelam tais diferenças.

Mapas físicos

Os **mapas físicos** baseiam-se na análise direta do DNA, e eles colocam os genes em relação às distâncias medidas em número de pares de bases, quilobases ou megabases. Um tipo comum de mapa físico é o que conecta segmentos isolados de DNA genômico que foi clonado nas bactérias ou leveduras (**Figura 20.3**). Em geral, os mapas físicos têm maior resolução e são mais exatos que os mapas genéticos. Um mapa físico é análogo a um mapa da vizinhança que mostra a localização de cada casa ao longo de uma rua, enquanto um mapa genético é análogo a um mapa de rodovia que mostra as localizações gerais das principais cidades e municípios.

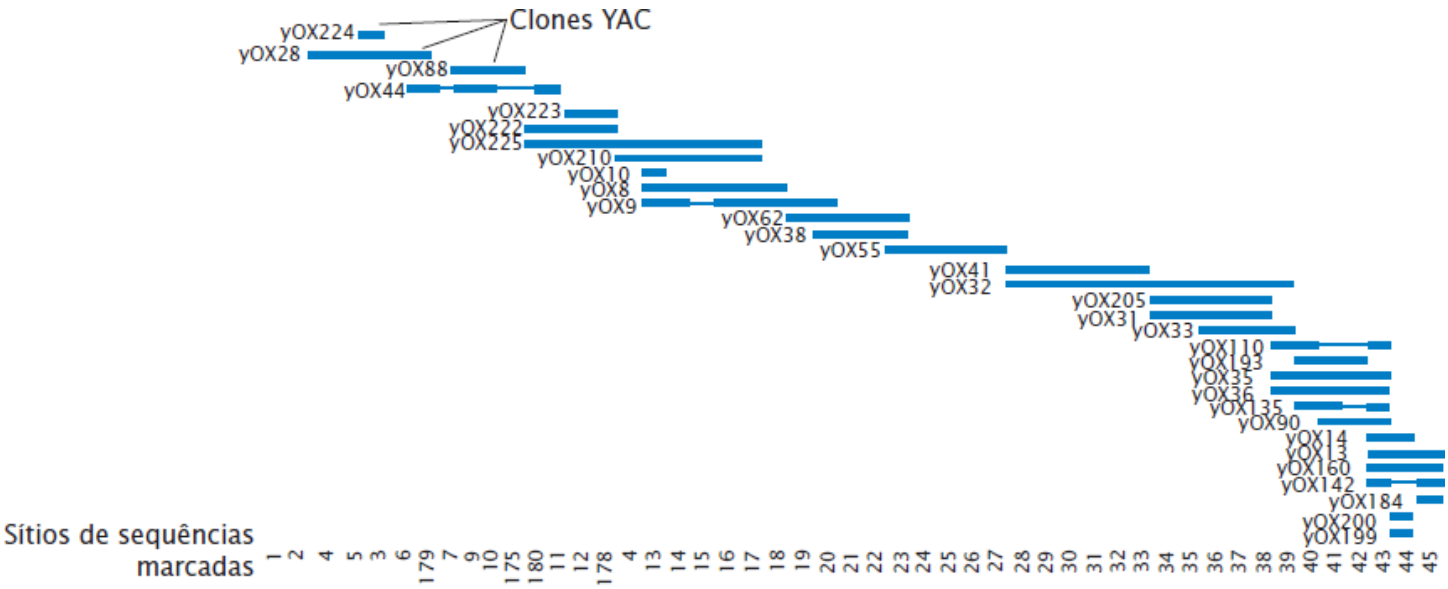


Figura 20.3 Os mapas físicos são, com frequência, usados com o propósito de ordenar os fragmentos de DNA clonados. Aqui temos uma parte de um mapa físico de um conjunto de clones YAC sobrepostos de uma extremidade do cromossomo Y humano.

Uma das técnicas que é usada para criar os mapas físicos é o mapeamento de restrição, que determina a posição dos sítios de restrição no DNA. Quando um segmento de DNA é cortado com uma enzima de restrição e os fragmentos são separados por eletroforese em gel, o número de sítios de restrição no DNA e as distâncias entre eles podem ser determinados pelo número e pelas posições das bandas no gel (ver *Capítulo 19*). Entretanto, esta informação não nos diz a ordem ou a localização exata dos sítios de restrição. Para mapear os sítios de restrição, uma amostra de DNA é cortada com uma enzima de restrição e outra amostra é cortada com uma enzima de restrição diferente. Uma terceira amostra é cortada com ambas as enzimas de restrição juntas (uma dupla digestão). Os fragmentos de DNA produzidos por estas digestões de restrição são separados por eletroforese em gel e seus tamanhos são comparados. A sobreposição no tamanho dos fragmentos produzidos pelas digestões pode ser usada para posicionar os sítios de restrição na molécula de DNA original (ver o Problema desenvolvido no final deste capítulo). Mais mapeamento de restrição é feito com várias enzimas de restrição, usadas sozinhas ou em várias combinações, produzindo muitos fragmentos de restrição. Com segmentos longos de DNA (maiores que 30 kb), programas de computador são usados para determinar os mapas de restrição e o mapeamento de restrição pode ser facilitado ao se marcar uma extremidade de um grande fragmento de DNA com radioatividade ou ao identificar a extremidade com o uso de uma sonda.

Os mapas físicos, como os mapas de restrição de fragmentos de DNA ou até de cromossomos inteiros, são criados para análise genômica. Estes mapas compridos são criados ao combinar os mapas de fragmentos genômicos curtos e sobrepostos. Existem várias técnicas adicionais para criar mapas físicos, incluindo: mapeamento de sítio de sequência marcada (STS), que localiza as posições de sequências curtas únicas de DNA em um cromossomo; hibridização *in situ*, na qual os marcadores podem ser mapeados visualmente nas localizações nos cromossomos; e sequenciamento de DNA.

Conceitos

Os mapas genéticos e físicos fornecem informações sobre as posições e distâncias relativas entre os genes, marcadores moleculares e segmentos de cromossomos. Os mapas genéticos baseiam-se nas taxas de recombinação e são medidos em porcentagem de recombinação (unidades de mapa) ou centimorgans. Os mapas físicos baseiam-se nas distâncias físicas e são medidos em pares de bases.

Checagem dos conceitos 1

Quais são algumas das limitações dos mapas genéticos?

Como sequenciar um genoma inteiro

Os primeiros genomas a serem sequenciados foram os de pequenos vírus. O genoma do bacteriófago λ, composto por 49.000 pb, foi completado em 1982. Em 1995, o primeiro genoma de um organismo de vida livre (*Haemophilus influenzae*) foi sequenciado por Craig Venter e Claire Fraser do Institute for Genomic Research (TIGR) e Hamilton Smith da Johns Hopkins University. Esta bactéria tem um pequeno genoma de 1,8 milhão de pares de base (**Figura 20.4**). Em 1996, foi determinado o genoma do primeiro organismo eucariótico (levedura), seguido dos genomas da *Escherichia coli* (1997), *Caenorhabditis elegans* (1998), *Drosophila melanogaster* (2000) e *Arabidopsis thaliana* (2000). O primeiro esboço do genoma humano foi

completado em junho de 2000.

O objetivo principal da genômica estrutural é determinar a sequência de nucleotídios ordenada do genoma inteiro dos organismos. No Capítulo 19, analisamos alguns dos métodos usados para sequenciar pequenos fragmentos de DNA. O principal obstáculo para sequenciar um genoma inteiro é o imenso tamanho da maioria destes. Os genomas bacterianos têm pelo menos vários milhões de pares de base de comprimento, muitos genomas eucarióticos têm bilhões de pares de bases e estão distribuídos entre dúzias de cromossomos. Além disso, por motivos técnicos, o sequenciamento não pode iniciar em uma extremidade de um cromossomo e continuar direto até a outra extremidade; apenas fragmentos pequenos de DNA, em geral entre 500 e 700 nucleotídios, podem ser sequenciados de uma vez. Portanto, o sequenciamento de um genoma inteiro exige que o DNA seja rompido em milhares ou milhões de fragmentos menores para então ser sequenciado. A dificuldade está em colocar estas sequências pequenas de volta na ordem correta. São usadas duas alternativas para montar os pequenos fragmentos sequenciados em um genoma completo: sequenciamento baseado no mapa e sequenciamento *shotgun* do genoma completo. Vamos analisar duas destas abordagens no contexto do Projeto Genoma Humano.

Projeto Genoma Humano

Em 1980, os métodos para mapeamento e sequenciamento de fragmentos de DNA estavam desenvolvidos o suficiente para que os geneticistas começassem a propor que o genoma humano inteiro poderia ser sequenciado. Foi planejada uma colaboração internacional para executar o Projeto Genoma Humano; as estimativas iniciais sugeriram que seriam necessários 15 anos e 3 bilhões de dólares para cumprir a tarefa.

Esse projeto começou oficialmente em outubro de 1990. Os esforços iniciais se voltavam para desenvolver novos métodos automáticos para clonar e sequenciar o DNA e gerar mapas físicos e genéticos detalhados do genoma humano. O esforço foi um projeto público composto pela colaboração internacional de 20 grupos de pesquisa e centenas de pesquisadores individuais que formaram o International Human Genome Sequencing Consortium. Este grupo usou uma estratégia baseada em mapa para sequenciar o genoma humano.

Sequenciamento baseado no mapa. No **sequenciamento baseado no mapa**, fragmentos curtos sequenciados são montados em uma sequência de genoma inteiro criando-se primeiro os mapas genéticos e físicos detalhados do genoma, que fornecem localizações conhecidas dos marcadores genéticos (sítios de restrição, outros genes ou sequências de DNA conhecidas) em intervalos regulares ao longo de cada cromossomo. Estes marcadores são usados posteriormente para ajudar a alinhar os fragmentos curtos sequenciados na sua ordem correta.

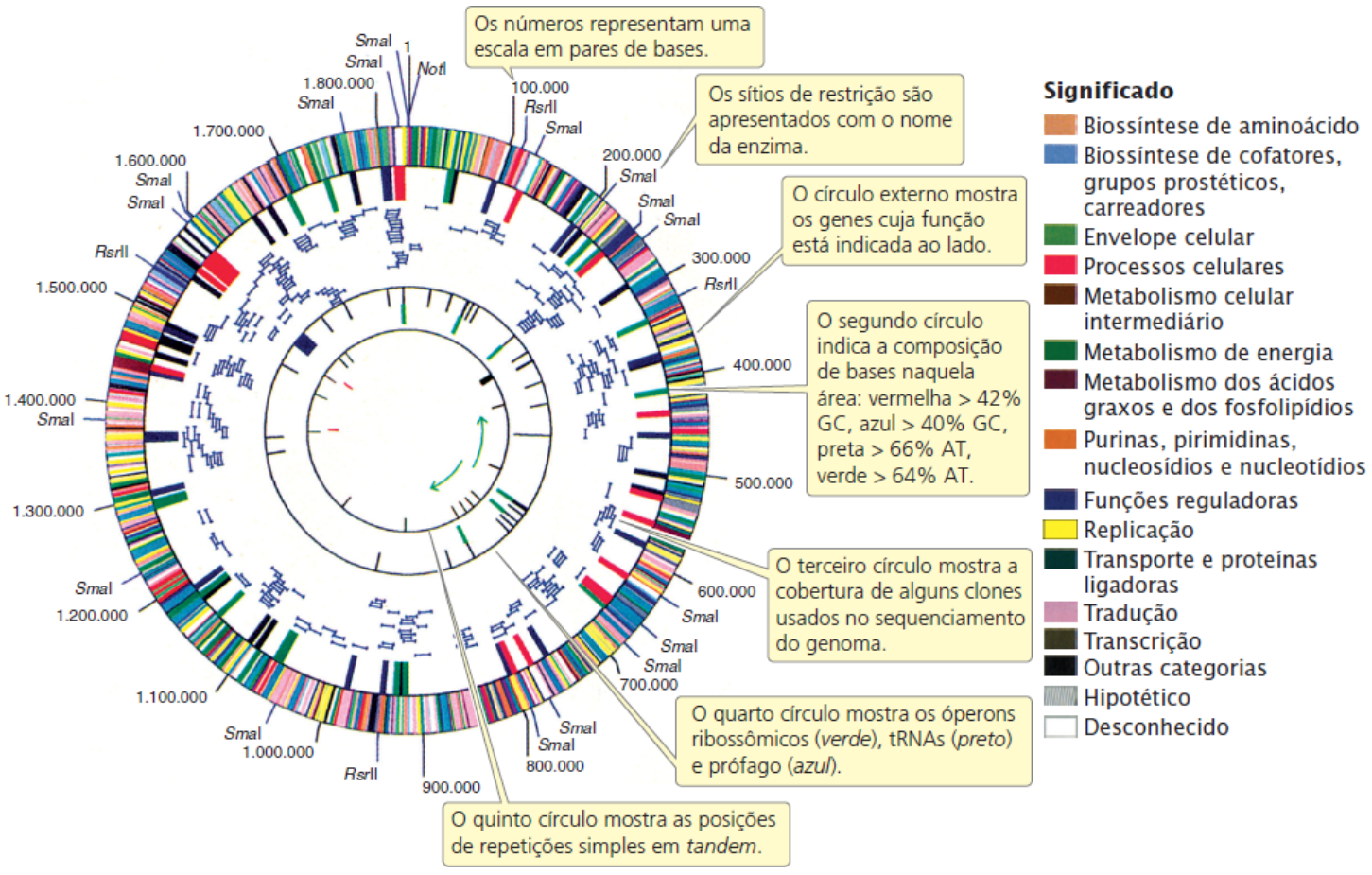
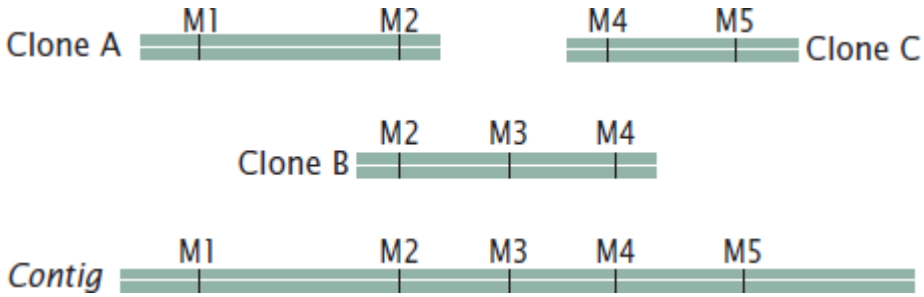


Figura 20.4 A bactéria *Haemophilas influenzae* foi o primeiro organismo de vida livre a ser sequenciado. (De R.D. Fleischman et al., Science 28 July 1995: 269 [5223], 496-512. Reimpressa com permissão de AAAS; cortesia da varredura de TIGR.)

Após os mapas genéticos e físicos estarem disponíveis, os cromossomos ou grandes segmentos de cromossomos são separados por eletroforese em gel de campo pulsado (PFGE) ou citometria de fluxo. A eletroforese em gel padrão não consegue separar segmentos muito grandes de DNA, como cromossomos inteiros, mas a PFGE consegue separar grandes moléculas de DNA ou cromossomos inteiros em um gel ao alternar periodicamente a orientação de um campo elétrico. Na citometria de fluxo, os cromossomos são classificados opticamente pelo tamanho.

Cada cromossomo (ou às vezes, o genoma inteiro) é então cortado por digestão parcial com enzimas de restrição (**Figura 20.5**). A digestão parcial significa que as enzimas de restrição atuam por um tempo limitado de modo que nem todos os sítios de restrição em cada molécula de DNA sejam cortados. A digestão parcial produz um conjunto de grandes fragmentos de DNA sobrepostos, que são clonados com o uso de cosmídios, cromossomos artificiais de levedura (YACS) ou cromossomos bacterianos artificiais (BACs, ver *Capítulo 19*).

A seguir, esses grandes clones de inserção são colocados na sua ordem correta no cromossomo (ver Figura 20.5). Esta montagem pode ser feita de várias formas. Um método depende da existência de um mapa de alta densidade dos marcadores genéticos. Uma sonda de DNA complementar é preparada para cada marcador genético e uma biblioteca de grandes clones de inserção é selecionada com uma sonda, que vai hibridizar com qualquer colônia contendo um clone com o marcador. A biblioteca é selecionada para os marcadores vizinhos. Como os clones são muito maiores que os marcadores usados como sondas, alguns clones têm mais de um marcador. Por exemplo, o clone A pode ter os marcadores M1 e M2, o clone M os marcadores M2, M3 e M4 e o clone C os marcadores M4 e M5. Tal resultado indicaria que estes clones têm áreas de sobreposição, como apresentado aqui:



Um conjunto de dois ou mais fragmentos de DNA sobrepostos que formam um segmento contínuo de DNA é chamado de ***contig***. Esta abordagem foi usada em 1993 para criar um *contig* com 196 clones YAC sobrepostos (ver Figura 20.3) do cromossomo Y humano.

A ordem dos clones também pode ser determinada sem o uso de mapas genéticos preexistentes. Por exemplo, cada clone pode ser cortado com várias enzimas de restrição e os fragmentos resultantes são separados por eletroforese em gel. Este método gera um conjunto único de fragmentos de restrição, chamado de impressão digital (*fingerprint*) para cada clone. Os padrões de restrição para os clones são armazenados em um banco de dados. Um programa de computador é usado para examinar os padrões de restrição de todos os clones e procurar áreas de sobreposição. A sobreposição é então usada para ordenar os clones, como apresentado aqui:

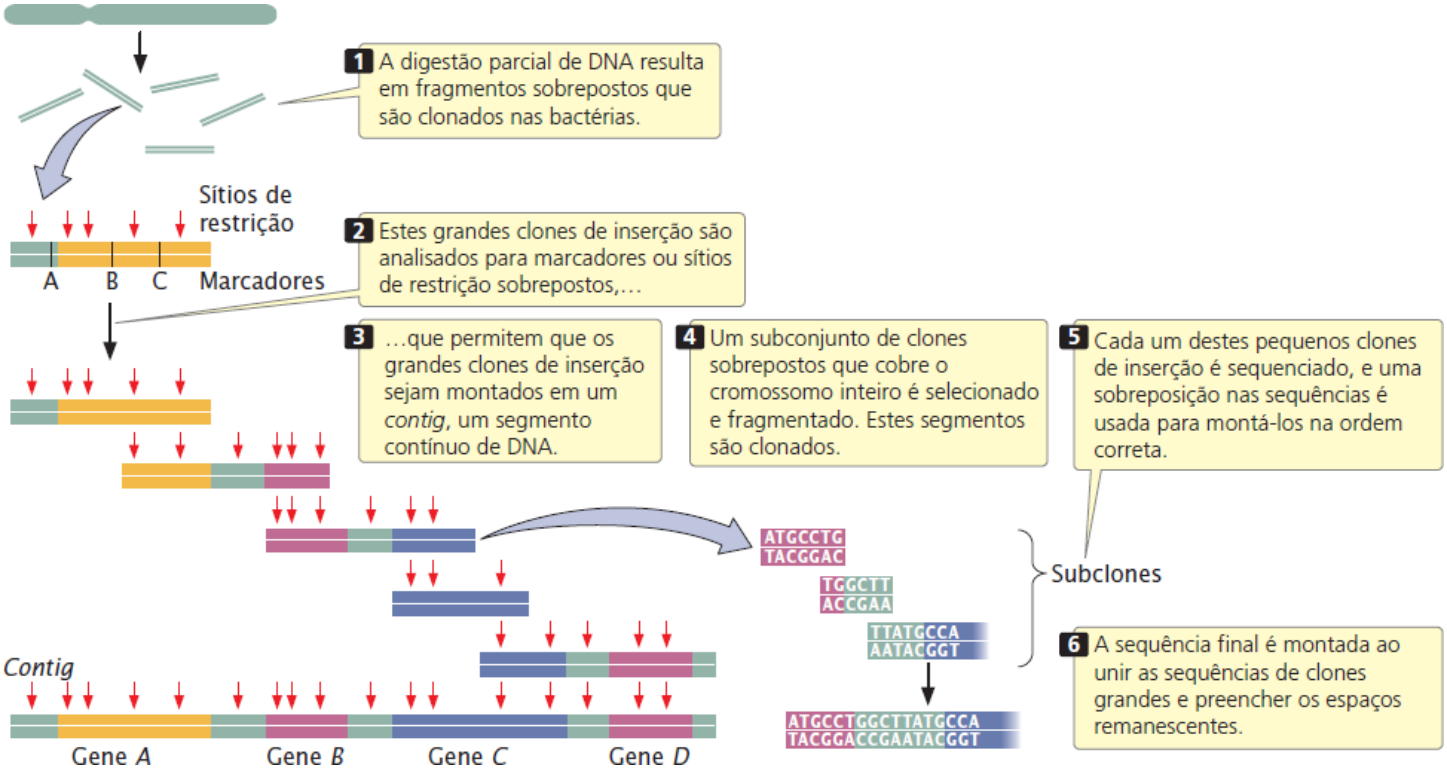
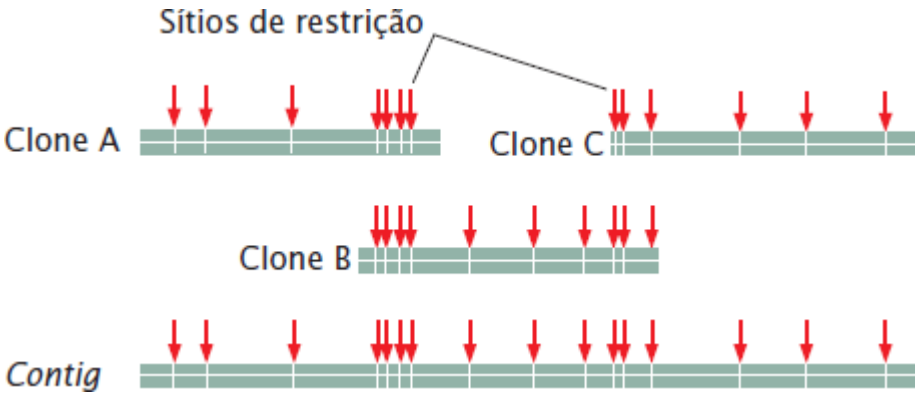


Figura 20.5 **As abordagens baseadas em mapa para o sequenciamento do genoma inteiro dependem de mapas genéticos e físicos detalhados para alinhar os fragmentos sequenciados.**



Outros marcadores genéticos podem ser usados para ajudar a posicionar os *contigs* ao longo do cromossomo. **Resolva o Problema 27**

Quando os grandes clones de inserção são montados na ordem correta no cromossomo, um subconjunto de clones sobrepostos que cobrem eficientemente o cromossomo inteiro pode ser selecionado para o sequenciamento; o objetivo é selecionar o número mínimo de clones necessários para representar o cromossomo. Cada um dos grandes clones de inserção selecionados é clivado em fragmentos sobrepostos menores que são clonados (ver Figura 20.5). Estes clones menores (chamados de pequenos clones de inserção) são sequenciados. As sequências dos pequenos clones de inserção são examinadas quanto à sobreposição, o que permite que sejam corretamente montadas para originar a sequência dos grandes clones de inserção. Uma quantidade suficiente de pequenos clones de inserção sobrepostos é sequenciada para garantir que o genoma inteiro seja sequenciado várias vezes. Finalmente, o genoma inteiro é montado ao colocar as sequências de todos os *contigs* sobrepostos (ver Figura 20.5). Mesmo assim, os espaços no mapa do genoma ainda existem e precisam ser preenchidos ao usar outros métodos.

O International Human Genome Squencing Consortium usou uma abordagem semelhante baseada no mapa para sequenciar o genoma humano. Muitas cópias do genoma humano foram cortadas em fragmentos de cerca de 150.000 pb cada, que foram inseridos em cromossomos bacterianos artificiais. Os cromossomos artificiais de levedura e cosmídios foram usados nos estágios iniciais do projeto, mas não eram estáveis como os clones BAC, embora os clones YAC tenham sido muito úteis na reunião de alguns dos *contigs* maiores. *Fingerprints* de restrição foram usados para montar os clones BAC nos *contigs*, que eram posicionados nos cromossomos com o uso de marcadores genéticos e sondas. Os clones BAC eram aparados em fragmentos sobrepostos menores e sequenciados e o genoma inteiro foi montado ao reunir a sequência dos clones BAC.

Sequenciamento shotgun do genoma completo. Em 1988, Craigh Venter anunciou que lideraria uma empresa

chamada Celera Genomics em um esforço direcionado para o sequenciamento do genoma humano. Ele propôs usar uma abordagem de sequenciamento pela estratégia *shotgun*, que seria mais rápida do que a abordagem baseada em mapa empregada pelo International Human Genome Sequencing Consortium. No **sequenciamento *shotgun* do genoma completo** (**Figura 20.6**), pequenos clones de inserção são preparados diretamente a partir do DNA genômico e sequenciados. Poderosos programas de computador montam o genoma inteiro por meio de exame da sobreposição dos pequenos clones de inserção. Uma vantagem deste tipo de sequenciamento é que os pequenos clones de inserção podem ser colocados em plasmídios, que são simples e fáceis de manipular. A exigência de sobreposição significa que a maior parte do genoma será sequenciada múltiplas vezes (de 10 a 15 vezes). O número médio de vezes que um nucleotídio no genoma é sequenciado é chamado de cobertura de sequenciamento. Por exemplo, cobertura 10 × significa que um nucleotídio médio no genoma foi sequenciado 10 vezes.

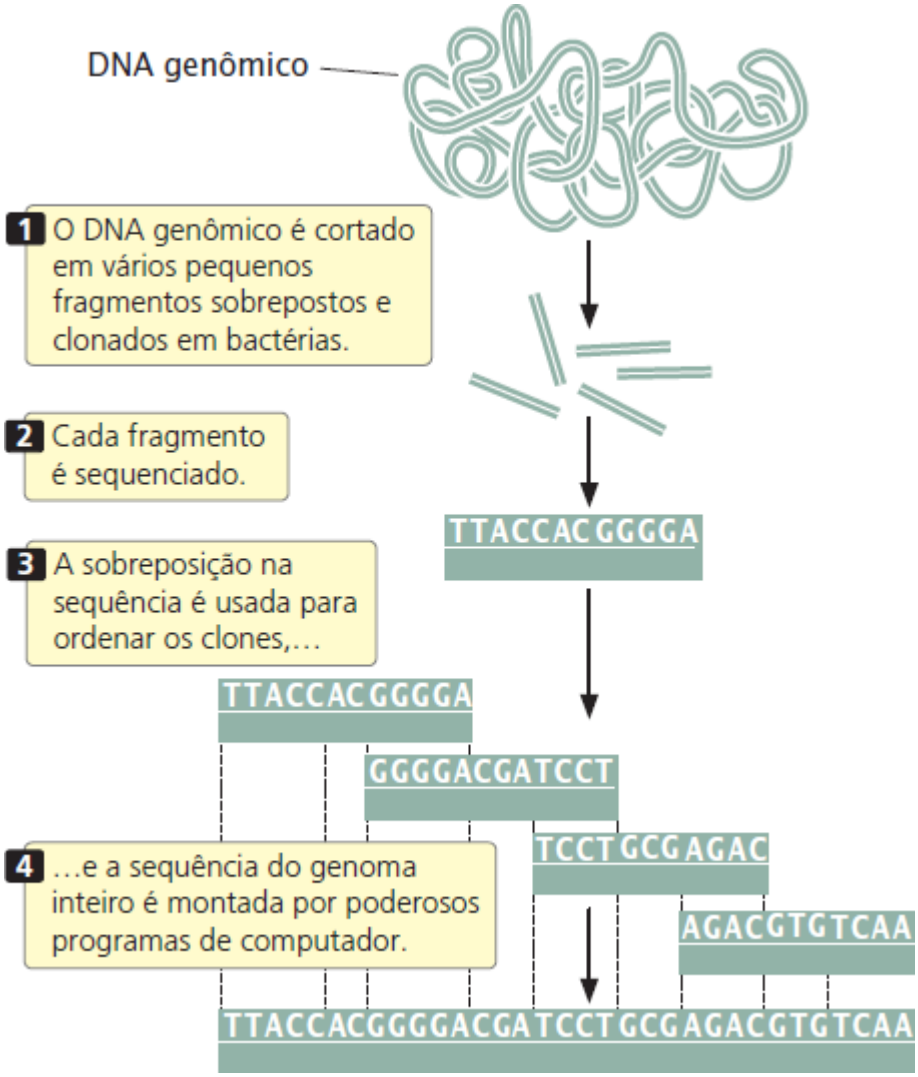


Figura 20.6 O sequenciamento *shotgun* do genoma completo usa a sobreposição de sequência para alinhar fragmentos sequenciados.

O sequenciamento *shotgun* foi usado inicialmente para montar pequenos genomas como os das bactérias. Quando Venter propôs o uso deste método para sequenciar o genoma humano, não estava muito claro se o projeto conseguiria montar com sucesso um genoma complexo composto por bilhões de pares de bases como o genoma humano. Hoje em dia, praticamente todos os genomas são sequenciados usando o método *shotgun* do genoma completo.

Conceitos

O sequenciamento de um genoma demanda a quebra dele em pequenos segmentos sobrepostos dos quais as sequências de DNA podem ser determinadas em uma reação de sequenciamento. No sequenciamento baseado no mapa, os fragmentos sequenciados são ordenados na sequência final do genoma com o uso de mapas genéticos e físicos. No sequenciamento *shotgun* do genoma completo, o genoma é montado pela comparação da sobreposição nas sequências dos pequenos fragmentos.

Checagem dos conceitos 2

Um *contig* é:

a. Um conjunto de marcadores moleculares usados no mapeamento genético.
b. Um conjunto de fragmentos sobrepostos que formam um intervalo contínuo de DNA.
c. Um conjunto de fragmentos gerados por uma enzima de restrição.
d. Um pequeno fragmento de DNA usado no sequenciamento.

Resultados e implicações do Projeto Genoma Humano. No verão de 2000, os projetos públicos e privados de sequenciamento anunciaram o término de um esboço preliminar que incluía a maior parte da sequência do genoma humano, 5 anos à frente do calendário. A sequência do genoma humano foi declarada completa na primavera de 2003, embora ainda existam algumas lacunas. Para a maioria dos cromossomos, a sequência final tem exatidão (acurácia) de 99,999%, com menos de um par de bases de erro por 100.000 pb, uma taxa de exatidão 10 vezes maior do que a meta inicial.

Com o primeiro genoma humano determinado, é muito mais fácil sequenciar genomas adicionais. Agora é possível sequenciar um genoma humano inteiro em um único dia e os genomas de milhares de pessoas foram sequenciados. O custo de sequenciar um genoma humano completo também caiu drasticamente e continuará a cair à medida que a tecnologia de sequenciamento avançar. Várias empresas estão investindo para desenvolver a tecnologia para sequenciar um genoma humano completo por US$1.000 dólares ou menos.

Foi comprovado que a disponibilidade da sequência completa do genoma humano é muito útil. A sequência forneceu ferramentas para detectar e mapear variantes genéticas no genoma humano, facilitando muito o mapeamento gênico nos humanos. Por exemplo, foram definidos muitos milhões de sítios onde as pessoas têm diferença em um único nucleotídio (chamados de polimorfismos de único nucleotídio, ver próxima seção), e estes sítios são usados em estudos de associação genômica ampla para localizar genes que afetam doenças e traços nos humanos. A sequência também fornece informações importantes sobre o desenvolvimento e muitos processos celulares básicos.

São usadas técnicas de sequenciamento da próxima geração que permitem o sequenciamento rápido e barato do DNA genômico (ver *Capítulo 19*) para abordar questões fundamentais em muitas áreas. Por exemplo, os genomas de várias células de câncer foram completamente sequenciados e comparados com sequências das células saudáveis da mesma pessoa, permitindo a determinação completa de todas as mutações que levam a formação do tumor e o progresso do câncer. O genoma completo de um recém-nascido foi recentemente sequenciado a partir do DNA fetal isolado do sangue de sua mãe. O 1000 Genomes Project está sequenciando e comparando os genomas de milhares de pessoas de diferentes grupos étnicos, com o objetivo de detectar a maioria das variações mais comuns que existem na espécie humana. O sequenciamento dos genomas completos de pais e de seus filhos permite uma estimativa direta das taxas de mutação.

O DNA foi extraído dos ossos de humanos ancestrais, incluindo neandertais e denisovanos (um pequeno grupo de humanos que parece estar intimamente relacionado com os neandertais) e completamente sequenciado. As comparações do genoma humano moderno com esta e outras espécies estão aumentando nossa compreensão sobre a evolução do ser humano, assim como nosso conhecimento sobre a evolução e história de toda vida.

Apesar destes benefícios e sucessos, algumas pessoas estão desapontadas com a falta de resultados tangíveis do Projeto Genoma Humano. Quando foi proposto, houve uma especulação que o sequenciamento do genoma humano causaria uma revolução imediata na prática da medicina, levando a nova compreensão para tratar doenças comuns e impulsionando o desenvolvimento de novos fármacos poderosos. Embora os dados da sequência genômica tenham produzidos numerosos novos e interessantes achados e tenham levado a melhor compreensão de muitas doenças, os dados ainda são raramente usados pelos médicos no tratamento dos pacientes. Sem dúvida, as informações sobre o genoma serão importantes para a medicina no futuro, tanto para adaptar o tratamento para pacientes individuais (medicina personalizada) quanto para descoberta de medicamentos, mas ainda não está certo quando esta mudança vai acontecer.

Junto com muitos potenciais benefícios de ter a informação da sequência completa, existem preocupações sobre seu uso indevido. Com o conhecimento obtido a partir do sequenciamento genômico, muito mais genes para doenças, distúrbios e traços comportamentais e físicos serão identificados, aumentando o número de testes genéticos que podem ser realizados para fazer previsões sobre o futuro do genótipo e da saúde de uma pessoa. Algumas pessoas estão preocupadas que as informações obtidas com exames genéticos possam ser usadas para discriminar entre as pessoas que são carreadoras de genes causadores de doenças ou que possam estar sob risco de desenvolver uma doença no futuro. Essa questão foi abordada nos EUA com a aprovação do *Genetic Information Nondiscrimination Act* (ver *Capítulo 6*), que proíbe que as empresas de seguro de saúde e empregadores usem informações genéticas para tomar decisões sobre a cobertura de seguro de saúde e emprego (embora não se aplique a vida, incapacidade e seguro de saúde a longo prazo). Também surgiram dúvidas sobre quem deve ter acesso à sequência genômica de uma pessoa. E os parentes, que tenham genomas semelhantes e possam estar sob risco de alguma das mesmas doenças? Também existem questões sobre o uso destas informações para selecionar traços específicos nos futuros descendentes. Todas estas questões são legítimas e devem ser abordadas se queremos usar as informações do sequenciamento do genoma com

responsabilidade.

Conceitos

O Projeto Genoma Humano foi um esforço para sequenciar o genoma humano inteiro. Duas equipes concorrentes completaram um esboço preliminar da sequência; ambas terminaram este esboço da sequência do genoma em 2000. O genoma inteiro foi sequenciado em 2003. A capacidade de sequenciar rapidamente os genomas humanos levanta várias questões éticas.

Polimorfismos de único nucleotídio

Desde o término do sequenciamento do genoma humano, os sequenciadores voltaram parte do seu esforço para as diferenças de mapeamento entre as pessoas nas suas sequências genômicas.

Imagine que você está subindo o elevador com um estranho qualquer. Quanto do seu genoma você tem em comum com esta pessoa? Estudos de variação no genoma humano indicam que você e o estranho serão idênticos em cerca de 99,9% de suas sequências de DNA. A diferença entre você e o estranho é muito pequena em termos *relativos*, mas como o genoma humano é tão grande (3,2 bilhões de pares de bases), você e o estranho serão diferentes em mais de 3 milhões de pares de bases do seu DNA genômico. Estas diferenças nos tornam únicos e afetam muito nossas características físicas, nossa saúde e até nossa inteligência e personalidade.

Um sítio no genoma onde os membros de uma espécie diferem em um único par de bases é chamado de **polimorfismo de único nucleotídio** (SNP, pronunciado “snip”). Os SNPs surgem por mutação e são herdados como variantes alélicas (como são os alelos que produzem diferenças fenotípicas, como os tipos sanguíneos), embora os SNPs não produzam uma diferença fenotípica. Os polimorfismos de único nucleotídio são numerosos e estão presentes por todo o genoma. Em uma comparação do mesmo cromossomo de duas pessoas diferentes, um SNP pode ser encontrado aproximadamente a cada 1.000 pb.

Haplótipos e desequilíbrio de ligação. A maioria dos SNPs presente dentro de uma população surge uma vez a partir de uma mutação simples que ocorre em um cromossomo específico e subsequentemente se espalhou pela população. Assim, cada SNP está inicialmente associado a outros SNPs (assim como outros tipos de variantes genéticas ou alelos) que existiam nesse cromossomo no qual surgiu a mutação. O conjunto específico de SNPs e outras variantes genéticas observados em um único cromossomo ou parte de um cromossomo é chamado de **haplótipo** (**Figura 20.7**). Os SNPs dentro de um haplótipo estão fisicamente ligados e, portanto, tendem a ser herdados juntos. Novos haplótipos podem surgir por meio da mutação ou do *crossing over*, que rompe o conjunto particular de SNPs em um haplótipo. Como a taxa de *crossing over* é proporcional às distâncias físicas entre os genes, SNPs e outras variantes genéticas próximas no cromossomo estarão fortemente associadas aos haplótipos. A associação não aleatória entre as variantes genéticas dentro de um haplótipo é chamada de **desequilíbrio de ligação**. Como os SNPs em um haplótipo são herdados juntos, um haplótipo composto por milhares de SNPs pode ser identificado com o uso de apenas alguns SNPs. Alguns SNPs usados para identificar um haplótipo são chamados de **SNPs marcados**. Existem cerca de 10 milhões de SNPs na população humana, mas por causa do desequilíbrio de ligação, estes SNPs estão presentes em um número muito menor de haplótipos. Portanto, um número relativamente pequeno de SNPs, talvez apenas 100.000, pode ser usado para identificar a maioria dos haplótipos nos humanos.

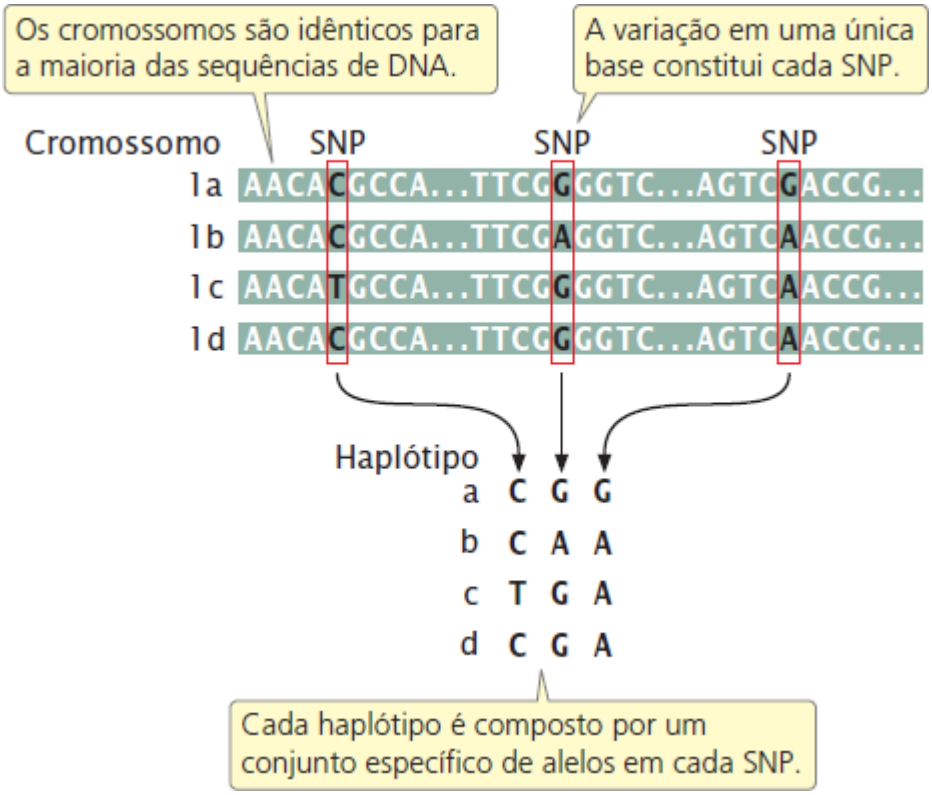


Figura 20.7 Um haplótipo é um conjunto específico de SNPs e outras variantes genéticas observadas em um único cromossomo ou parte de um cromossomo. Os cromossomos 1a, 1b, 1c e 1d representam diferentes cópias de um cromossomo que pode ser encontrado em uma população.

Uso de polimorfismos de único nucleotídio. Graças a sua variabilidade e ocorrência disseminada pelo genoma, os SNPs são marcadores valiosos nos estudos de ligação. Quando um SNP está fisicamente próximo de um *locus* causador de doença, ele tenderá a ser herdado junto com o alelo responsável pela doença. Pessoas com a doença tenderão a ter SNPs diferentes em relação às pessoas saudáveis. Uma comparação dos haplótipos de SNP em pessoas com uma doença e nas pessoas saudáveis pode revelar a presença de genes que afetam a doença; como o gene da doença e o SNP estão intimamente ligados, a localização do gene causador da doença pode ser determinado a partir da localização dos SNPs associados. Esta abordagem é a mesma que a usada no mapeamento gênico com RFLPs (ver *Capítulo 19*), mas existem mais SNPs do que RFLPs, fornecendo um denso conjunto de marcadores variáveis que cobrem o genoma inteiro que podem ser usados mais efetivamente no mapeamento.

Um esforço internacional chamado o Projeto HapMap Internacional começou em 2002 com o objetivo de catalogar e mapear SNPs e outras variantes genômicas que podem ser usadas para identificar haplótipos comuns nas populações humanas para usar em estudos de ligação e família. A Fase I do projeto, completada em 2005, catalogou mais de 1 milhão de SNPs nos genomas de 269 pessoas para quatro diferentes populações humanas (africana, japonesa, chinesa e europeia). Esses SNPs estão espalhados em todos os 23 cromossomos em uma distância de um SNP por 5.000 pb. A Fase II do projeto, completada em 2006, catalogou um total de 4,6 milhões de SNPs.

Os alelos a partir da maioria dos SNPs comuns são encontrados em todos os quatro grupos étnicos, embora as frequências dos alelos variem muito entre as populações. A maior diversidade genética de SNPs é encontrada dentro dos africanos, o que é consistente com muitos outros estudos que sugerem que os humanos evoluíram primeiro na África.

Dados do Projeto HapMap forneceram informações importantes sobre a função e evolução do genoma humano. Por exemplo, estudos de desequilíbrio de ligação com o uso de SNPs determinaram que a recombinação não ocorre aleatoriamente entre os cromossomos: existem numerosos *hotspots* de recombinação, onde mais recombinação ocorre do que simplesmente o esperado pelo acaso. A distribuição de SNPs também fornece informações sobre as regiões do genoma humano que evoluíram rapidamente no passado recente, fornecendo dicas sobre como os humanos responderam à seleção natural.

Estudos de associação genômica ampla. Muitas doenças comuns são causadas por interações complexas de múltiplos genes: a disponibilidade de SNPs facilita muito a pesquisa por estes genes. Os **estudos de associação genômica ampla** usam numerosos SNPs dispersos pelo genoma para encontrar os genes de interesse. Logo após terminar a fase I do Projeto HapMap, os pesquisadores usaram os dados do SNP para conduzir um estudo de associação genômica ampla quanto à presença de um gene principal que provoca a degeneração macular relacionada com a idade, uma das causas principais de cegueira entre os idosos. Neste estudo, os pesquisadores estudaram o genótipo de 96 pessoas com a degeneração macular e 50 pessoas com olhos saudáveis para mais de 100.000 SNPs que cobriam o genoma. O estudo revelou uma forte associação entre a doença e um gene

no cromossomo 1 que codifica o fator H de complemento, que é importante na função imunológica. Este achado sugere que a degeneração macular pode ser tratada com fármacos que afetem as proteínas do complemento. Outro estudo revelou um gene importante associado à doença de Crohn, um distúrbio inflamatório comum do trato gastrintestinal. Outros estudos usando varreduras de genoma amplo de SNPs identificaram genes que contribuem para cardiopatia, densidade óssea, câncer de próstata e diabetes.

Em uma aplicação bem-sucedida dos SNPs para encontrar associações das doenças, pesquisadores mapearam o genótipo de 17.000 pessoas no Reino Unido para 500.000 SNPs em 2007. Eles detectaram fortes associações entre 24 genes e segmentos de cromossomos e a incidência de sete doenças comuns, incluindo coronariopatia, doença de Crohn, artrite reumatoide, transtorno bipolar, hipertensão e dois tipos de diabetes. A importância deste estudo é sua demonstração que os estudos de associação genômica ampla utilizando SNPs poderiam localizar genes que contribuem para doenças complexas causadas por múltiplos fatores genéticos e ambientais.

Nos últimos anos, os SNPs foram usados nos estudos de associação genômica ampla para localizar genes que influenciam muitos traços adicionais, como a idade da puberdade e da menopausa nas mulheres, variação nas características faciais, pigmentação da pele, cor dos olhos, peso corporal, densidade óssea, glaucoma e até suscetibilidade a doenças infecciosas como a doença meningocócica e tuberculose. Infelizmente, os genes identificados explicam apenas uma pequena proporção da influência genética sobre o traço. Por exemplo, um imenso estudo de associação genômica ampla combinou dados de 100.000 humanos em uma tentativa para localizar os genes que codificam proteínas associadas à presença dos lipídios do sangue e cardiopatia. Embora o estudo tenha identificado 95 *loci* diferentes associados a traços de lipídios, estes genes correspondem a apenas 25 a 30% da variação genética total nestes traços.

As sequências de DNA que codificam a maior parte da variação genética ausente em tais traços, às vezes chamadas de “matéria escura do genoma”, permanecem, até o momento, indetectadas em sua grande maioria. A baixa porcentagem de variação explicada pela maioria dos estudos de associação genômica ampla significa que os genes identificados não são, por si sós, indicadores úteis do risco de herdar a doença ou traço. Contudo, a identificação de genes específicos que influenciam uma doença ou traço pode levar a uma melhor compreensão dos processos biológicos que produzem o fenótipo.

Variações no número de cópias

Uma pessoa diploide normalmente tem duas cópias de cada gene, um herdado da mãe e um do pai. Contudo, estudos do genoma humano revelaram diferenças entre as pessoas no número de cópias de grandes sequências de DNA (maiores que 1.000 pb); estas variações são chamadas de **variações no número de cópias** (CNVs). As variações no número de cópias podem incluir deleções, fazendo com que algumas pessoas tenham apenas uma única cópia de uma sequência, ou duplicações, de modo que algumas pessoas tenham mais de duas cópias. Um estudo das CNVs em 270 pessoas de quatro populações (as estudadas no Projeto HapMap) identificou um número surpreendente destes tipos de variantes: mais de 1.447 regiões genômicas variavam no número de cópias, incluindo 12% do genoma humano. Mas apenas uma pequena parte do genoma humano foi avaliada neste estudo, e as CNVs detectadas são apenas um pequeno subconjunto de tudo que existe. Muitas das CNVs incluíam grandes regiões de sequência de DNA, frequentemente muitas centenas de milhares de pares de base de comprimento. Assim, as pessoas apresentam diferenças não apenas em milhões de SNPs individuais, mas também no número de cópias de muitos segmentos maiores do genoma.

A maioria das CNVs tem múltiplos genes e afeta potencialmente o fenótipo ao alterar a dose gênica e a posição das sequências, que podem afetar a regulação dos genes próximos. De fato, vários estudos encontraram associações entre CNVs e doenças, e até entre CNVs e a variabilidade fenotípica normal nas populações humanas. Por exemplo, variações no número de cópias do gene *UGT2B17* contribuem para as diferenças no metabolismo da testosterona entre os indivíduos e afetam o risco de câncer de próstata. As variações no número de cópias estão associadas a doença de Crohn, artrite reumatoide, psoríase, esquizofrenia, autismo, diabetes melito e déficit intelectual.

Sítios de sequências marcadas e etiquetas de sequência expressa

Um **sítio de sequências marcadas** (STS) é uma sequência de DNA curta (200 a 500 pb), encontrada apenas uma vez no genoma, cuja localização cromossômica foi determinada. Os STSs são usados para determinar a localização genômica de um clone de DNA. Os pesquisadores desenvolveram *primers* de PCR para cada STS, permitindo a amplificação de STS por PCR. Desta forma, a PCR pode ser usada para detectar um STS em um conjunto de clones. Como a localização cromossômica do STS não é conhecida, o achado do STS em um determinado clone fornece informações sobre a origem no genoma desse clone.

Outro tipo de sequência identificada é uma **etiqueta de sequência expressa** (EST). Estas são sequências que correspondem ao DNA que é transcrito em mRNA. Na maioria dos organismos eucarióticos, apenas uma pequena porcentagem de DNA codifica as proteínas, por exemplo, nos humanos, cerca de 1,5% do DNA codifica proteínas. Se apenas os genes que codificam proteínas são o alvo, um exame do mRNA em vez da sequência genômica do DNA inteiro é mais eficiente. As etiquetas de

sequência expressa são obtidas ao isolar o RNA de uma célula e sujeitá-lo a transcrição reversa, produzindo um conjunto de fragmentos de cDNA que correspondem a moléculas de RNA da célula. Intervalos curtos a partir das extremidades destes fragmentos de cDNA são sequenciados, e a sequência obtida (chamada de etiqueta) fornece um marcador que identifica o fragmento de DNA. As etiquetas de sequência expressa podem ser usadas para encontrar genes ativos em um determinado tecido ou em um ponto específico no desenvolvimento.

Conceitos

Além de coletar os dados da sequência genômica, os projetos de genômica estão coletando bancos de dados de nucleotídios que variam entre os organismos individuais (polimorfismos de único nucleotídio, SNPs), variações no número de cópias de sequências (variações no número de cópias), sequências únicas cuja localização cromossômica foi mapeada (sítios de sequências marcadas, STSs) e marcadores associados a sequências transcritas (etiquetas de sequência expressa, ESTs).

Checagem dos conceitos 3

Qual foi o objetivo do projeto HapMap?

Bioinformática

Atualmente sequências de genoma completo já foram determinadas para numerosos organismos, com muitos projetos adicionais em andamento. Estes estudos estão produzindo uma quantidade incrível de dados de sequência. Catalogar, armazenar, recuperar e analisar este imenso conjunto de dados são os principais desafios da genética moderna. A **bioinformática** é um campo emergente composto por biologia molecular e ciência computacional que se concentra no desenvolvimento de bancos de dados, algoritmos de pesquisa em computador, programa de previsão gênica e outras ferramentas analíticas usadas para dar sentido aos dados de sequência de DNA, RNA e proteínas. A bioinformática desenvolve e aplica estas ferramentas para “explorar os dados”, extraindo informações úteis a partir de projetos de sequenciamento. O desenvolvimento e o uso de algoritmos e programas de computador para analisar DNA e os dados de sequências de proteínas ajudam a tornar a biologia molecular um campo mais quantitativo. Dados de sequência disponíveis em bancos de dados *on-line* possibilitam que cientistas e estudantes de todo mundo tenham acesso a este incrível recurso.

Vários bancos de dados foram estabelecidos para a coleta e a análise das informações de sequência de DNA e proteína. Os bancos de dados primários têm as informações da sequência, junto com informações que descrevem a fonte da sequência e sua determinação. Os bancos de dados secundários têm os resultados das análises feitas nos dados de sequência primários, como as informações sobre padrões de sequência específicos, variações, mutações e relações evolutivas.

Após um genoma ser sequenciado, uma das primeiras tarefas é identificar genes potenciais na sequência. Infelizmente, não existem características universais que marquem o início e o fim de um gene. A abundância de DNA em um genoma e as complexidades da estrutura do gene dificultam o achado de genes na sequência. Foram desenvolvidos programas de computador para examinar sequências específicas no DNA associadas a alguns genes. Existem duas alternativas gerais para encontrar os genes. A técnica *ab initio* examina a sequência à procura de características que geralmente são intragênicas. Por exemplo, genes codificadores de proteínas são caracterizados por um quadro de leitura aberto, que inclui um códon de início e um códon de parada no mesmo quadro de leitura. Sequências específicas marcam os locais de recomposição no início e na extremidade dos íntrons; outras sequências específicas estão presentes nos promotores imediatamente *upstream* dos códons de início. A abordagem comparativa procura por similaridade entre uma nova sequência e as sequências de todos os genes conhecidos. Se for encontrada uma combinação, então a nova sequência será considerada como sendo um gene semelhante. Alguns destes programas de computador conseguem examinar bancos de dados de EST e sequências de proteínas para ver se existe indício de que um gene potencial é expresso.

É importante observar que os programas que foram desenvolvidos para identificar genes com base na sequência de DNA não são perfeitos. Portanto, os números de genes relatados na maioria dos projetos de genoma são estimativas. A presença de múltiplos íntrons, recomposição alternativa, múltiplas cópias de alguns genes e parte do DNA não codificado entre os genes dificulta a identificação e contagem exata dos genes.

Após um gene ser identificado, ele tem de passar por **notação**, o que significa conectar as informações da sua sequência a outras informações sobre sua função e expressão, a proteína que codifica e informações sobre genes semelhantes em outras espécies. Existem vários métodos para sondar a função de um gene, que serão discutidos na Seção 20.2 sobre genômica funcional. Existem programas de computador para determinar se já foram encontradas sequências semelhantes, na mesma espécie ou em espécie diferente. O programa mais usado destes é o Basic Local Alignment Search Tool (BLAST). Para fazer uma pesquisa do BLAST, um pesquisador submete uma sequência e o programa pesquisa o banco de dados para qualquer

sequência com regiões de alta semelhança com a sequência em questão. O programa retorna todas as sequências no banco de dados que são semelhantes, junto com informações sobre o grau de semelhança e a importância da combinação (qual probabilidade de que a semelhança ocorra por acaso sozinha).

Bancos de dados adicionais têm informações sobre a diversidade na sequência, como e onde um genoma varia entre os organismos individuais. O genoma humano foi mapeado para milhões de SNPs e estão sendo coletadas informações sobre a frequência destas variantes em diferentes populações. Foram desenvolvidos bancos de dados adicionais para catalogar todas as mutações conhecidas por causar determinadas doenças, o Human Variome Project é o esforço para coletar e disponibilizar todas as variações genéticas que afetam a saúde humana. Também estão sendo compiladas informações importantes sobre os padrões de expressão de milhares de genes encontrados nos genomas. **Resolva o Problema 29**

Conceitos

A bioinformática é um campo interdisciplinar que combina biologia molecular e ciência computacional. Ela desenvolve bancos de dados de DNA e sequências de proteínas e ferramentas para analisar tais sequências.

Checagem dos conceitos 4

A técnica *ab initio* encontra os genes ao procurar por:

a. Sequências comuns encontradas na maioria dos genes.
b. Semelhança na sequência de genes conhecidos.
c. mRNA com o uso de hibridização *in situ*.
d. Fenótipos mutantes.

Metagenômica

Os avanços na tecnologia de sequenciamento, que o tornaram mais rápido e menos caro, agora fornecem a possibilidade de sequenciar não apenas os genomas de espécie individual, mas os genomas de comunidades inteiras de organismos. A **metagenômica** é um campo emergente de tecnologia de sequenciamento no qual as sequências de genoma de um grupo de organismos que habitam um ambiente comum são separadas e determinadas.

Até o momento, a metagenômica é aplicada para comunidades microbianas. Tradicionalmente, as bactérias são estudadas ao cultivar e analisá-las em laboratório. Entretanto, muitas bactérias não podem ser cultivadas com o uso de técnicas laboratoriais. A metagenômica analisa as comunidades microbianas ao extrair o DNA do ambiente, determinando suas sequências e reconstruindo a composição e função da comunidade com base nas sequências do genoma. Esta técnica permite a identificação e análise genética de espécies que não podem crescer no laboratório e nunca foram estudadas pelos métodos microbiológicos tradicionais. Os genomas inteiros de algumas espécies dominantes foram reconstruídos a partir de amostras do ambiente, fornecendo muitas informações sobre a biologia destes micróbios aos cientistas.

Um estudo metagenômico inicial analisou a comunidade microbiana encontrada na drenagem ácida de uma jazida e determinou que esta comunidade era composta apenas por algumas espécies bacterianas. Outro estudo, chamado de Global Ocean Sampling Expedition, seguiu a rota da viagem de Darwin no H.M.S. *Beagle* nos anos de 1800. Os cientistas coletaram amostras de oceano e usaram os métodos metagenômicos para determinar suas comunidades microbianas. Nesses estudos, os cientistas catalogaram sequências para mais de 6 milhões de proteínas, incluindo mais de 1.700 novas famílias de proteínas. Surgiram alguns resultados importantes a partir dos estudos metagenômicos. As análises de genomas bacterianos encontrados nas amostras do oceano levaram à descoberta de proteínas proteorrodopsinas, que são bombas de próton acionadas pela luz. A pesquisa posterior demonstrou que estas proteínas são encontradas em diversos grupos microbianos e são uma fonte principal de fluxo de energia nos oceanos do mundo.

Outros estudos metagenômicos examinaram os genes das bactérias que habitam o trato intestinal humano. Estas bactérias, junto com as que habitam a pele e outras partes do corpo, são chamadas de **microbioma** humano. O microbioma de uma pessoa inclui mais de 100 trilhões de células, mais de 10 vezes as células humanas, e têm 100 vezes mais genes que o genoma humano.

A pesquisa está demonstrando que a flora é importante para a saúde humana. Um estudo examinou a flora do intestino de pessoas obesas e magras. Dois grupos de bactérias são comuns no intestino humano. Bacteroidetes e Firmicutes. Os pesquisadores descobriram que as pessoas obesas tinham relativamente mais Firmicutes do que as pessoas magras e que a proporção de Firmicutes reduzia nas pessoas obesas que perdiam peso com dieta de baixas calorias. Os mesmos resultados foram observados nos camundongos obesos e magros. Em um experimento interessante, os pesquisadores transferiram as bactérias dos camundongos obesos para os camundongos magros. Os camundongos que receberam as bactérias dos camundongos obesos extraíam mais calorias do seu alimento e armazenavam mais gordura, sugerindo que a flora do intestino é importante para a obesidade.

As comunidades ecológicas são caracterizadas com base nas suas espécies. A metagenômica permite caracterizar as comunidades com base nos genes, uma abordagem chamada genecêntrica. Uma visão genecêntrica levanta novos questionamentos: Alguns tipos de genes são mais comuns em algumas comunidades do que em outras? Os genes são essenciais para o fluxo de energia e reciclagem de nutrientes dentro de uma comunidade? Por causa do tamanho dos seus genomas, as comunidades eucarióticas ainda não são o foco destes estudos, mas muitos pesquisadores esperam incluí-las no futuro.

Conceitos

Os estudos metagenômicos examinam os genomas das comunidades de organismos que habitam ambiente comum. Esta abordagem tem sido aplicada a comunidades microbianas e permite que a composição e constituição genética da comunidade seja determinada sem cultivar e isolar espécies individuais. Os estudos metagenômicos são uma fonte de novos conhecimentos importantes nas comunidades microbianas.

Biologia sintética

A capacidade de sequenciar e estudar genomas inteiros, junto com maior compreensão das informações genéticas necessárias para os processos biológicos básicos, agora fornecem a capacidade de criar, inteiramente a partir do zero, organismos que nunca existiram. A biologia sintética é um campo novo que procura planejar organismos que possam fornecer funções úteis, como os micróbios que fornecem energia limpa ou degradam resíduos tóxicos.

Os biólogos desta área já misturaram e combinaram partes de diferentes organismos para sintetizar os micróbios. Em 2002, os geneticistas recriaram o poliovírus ao unir segmentos de DNA que eram sintetizados no laboratório. O mais incrível, em 2010, David Gibson e seus colegas sintetizaram, a partir do zero, o genoma completo de 1,08 milhão de pares de bases da bactéria *Mycoplasma mycoides*. Eles iniciaram com milhares de segmentos de DNA que foram sintetizados no laboratório e então reunidos em segmentos sucessivamente maiores até terem montado uma cópia completa do genoma. Dentro do genoma sintético eles incluíram um conjunto de sequências de DNA que soletra, em código, um endereço de *e-mail*, os nomes dos pesquisadores que participaram do projeto e várias citações conhecidas. Finalmente, os pesquisadores transplantaram o genoma artificial para uma célula de uma espécie bacteriana diferente, *M. capricolum*, cujo genoma original foi removido. A nova célula começou a expressar os traços especificados pelo genoma sintético.

A biologia sintética também se estende para as células eucarióticas. Em 2011, os geneticistas substituíram 90.000 pb de DNA do cromossomo 9 da levedura e 30.000 pb do cromossomo 6 de levedura por DNA sintético. Estas células de levedura parcialmente sintéticas cresceram normalmente e exibiram apenas pequenas diferenças na expressão gênica. O objetivo principal destes experimentos é substituir o genoma inteiro de 12 milhões de pb da levedura por sequências humanas projetadas.

Estes tipos de experimentos também levantam vários questionamentos. A capacidade de novos genomas sob medida e misturar e combinar partes de diferentes organismos cria o potencial para a síntese de micróbios perigosos que podem criar caos ecológico se escaparem do laboratório ou poderiam ser usados em guerra biológica ou bioterrorismo. As discussões correntes entre os geneticistas, profissionais da área de ética, especialistas em segurança e políticos estão abordando estas preocupações e se os genomas sintéticos podem ser produzidos e usados com segurança.

20.2 A genômica funcional determina a função dos genes ao usar abordagens com base na genômica

Uma sequência genômica, por si só, tem uso limitado. Simplesmente conhecer uma sequência seria como ter um grande conjunto de enciclopédias sem ser capaz de ler: você reconheceria as diferentes letras, mas o texto não teria sentido. A **genômica funcional** caracteriza o que as sequências fazem, sua função. Os objetivos da genômica funcional incluem a identificação de todas as moléculas de RNA transcritas a partir de um genoma, chamado de **transcriptoma** daquele genoma e todas as proteínas codificadas pelo genoma, chamadas de **proteoma**. A genômica funcional usa a bioinformática e as abordagens experimentais laboratoriais na sua pesquisa para definir a função das sequências de DNA.

O Capítulo 19 analisou os vários métodos para identificar os genes e avaliar suas funções, incluindo a hibridização *in situ*, mutagênese experimental e o uso de animais transgênicos e *knockouts*. Estes métodos podem ser aplicados a genes individuais e fornecer informações importantes sobre as localizações e funções das informações genéticas. Nesta seção, vamos nos concentrar principalmente nos métodos que dependem de conhecer as sequências de outros genes ou os métodos que podem ser aplicados a vários genes simultaneamente.

Como prever a função a partir da sequência

A sequência de nucleotídios de um gene pode ser usada para prever a sequência de aminoácidos da proteína que ela codifica. A proteína pode então ser sintetizada ou isolada e suas propriedades estudadas para determinar sua função. Entretanto, esta análise bioquímica para compreender a função do gene consome tempo e é cara. O objetivo principal da genômica funcional tem sido desenvolver métodos computacionais que permitam que a função gênica seja identificada a partir da sequência de DNA sozinha, superando o processo trabalhoso de isolar e caracterizar as proteínas individuais.

Pesquisas de homologia. Um método computacional (muitas vezes o primeiro a ser empregado) para determinar a função gênica é fazer uma pesquisa de homologia, que depende de comparações das sequências de DNA e das proteínas do mesmo organismo e de diferentes organismos. Os genes com uma evolução próxima são chamados de **homólogos**. Os genes homólogos encontrados em diferentes espécies que evoluíram a partir do mesmo gene em um ancestral comum são chamados de **ortólogos** (**Figura 20.8)**. Por exemplo, os genomas dos camundongos e dos humanos têm um gene que codifica a subunidade alfa da hemoglobina; os genes da alfa-hemoglobina de camundongo e humana são chamados de ortólogos, porque ambos evoluíram a partir de um gene da alfa-hemoglobina de um ancestral mamífero comum. Os genes homólogos no mesmo organismo (que surgiram a partir da duplicação de um único gene no passado evolutivo) são chamados **parálogos** (ver Figura 20.8). No genoma humano existe um gene que codifica a subunidade alfa da hemoglobina e outro gene homólogo que codifica a subunidade beta da hemoglobina. Estes parálogos surgiram porque um gene ancestral sofreu duplicação e os dois genes resultantes divergiram durante a evolução, dando origem aos genes das subunidades alfa e beta (ver Figura 26.18). Os genes homólogos (ortólogos e parálogos) têm, frequentemente, as mesmas funções ou funções parecidas, então, após uma função ter sido designada para um determinado gene, esta constitui um indício da função de um gene homólogo.

Existem bancos de dados com genes e proteínas encontrados em uma gama de organismos para pesquisas de homologia. Foram desenvolvidos programas de computador poderosos, como BLAST, para procurar sequências específicas nesses bancos de dados. Suponha que um geneticista sequencie um genoma e localize um gene que codifica uma proteína de função desconhecida. Uma pesquisa de homologia conduzida em bancos de dados com sequências de DNA ou proteínas de outros organismos pode identificar uma ou mais sequências ortólogas. Se uma função for conhecida para uma proteína codificada por uma destas sequências, esta poderia fornecer informações sobre a função da proteína recém-descoberta. Às vezes, as pesquisas não identificam os genes homólogos, neste caso, os pesquisadores podem procurar por partes de um gene que são semelhantes a partes de outros genes com informação conhecida. **Resolva o Problema 33**

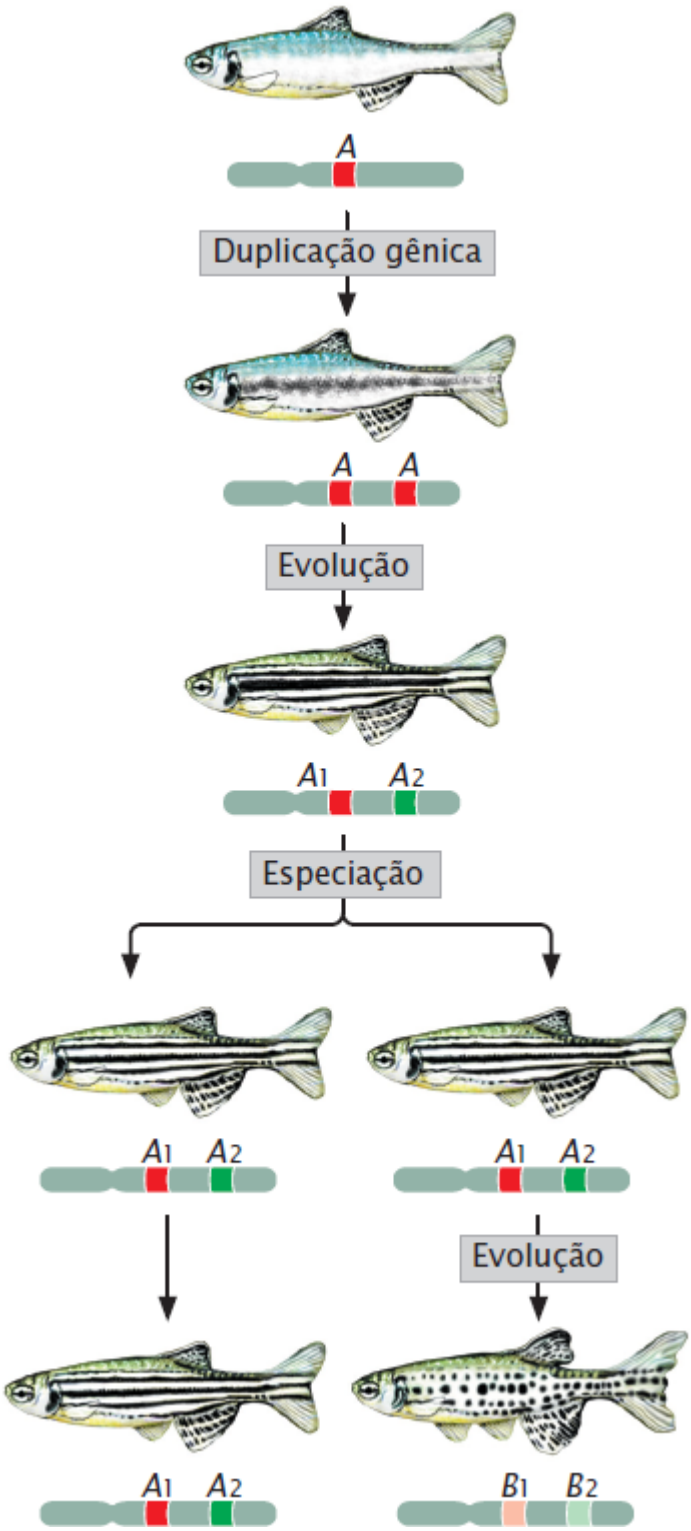


Figura 20.8 Os genes homólogos têm a mesma linha de evolução. Os ortólogos são os genes homólogos encontrados em diferentes espécies, enquanto os parálogos são genes homólogos encontrados na mesma espécie.

Domínios da proteína. As proteínas complexas têm, com frequência, regiões com formatos ou funções específicas chamadas de **domínios da proteína**. Por exemplo, determinadas proteínas ligadoras de DNA se conectam ao DNA da mesma forma; essas proteínas têm um domínio em comum que desempenha a função de ligação ao DNA. Cada proteína tem uma disposição de aminoácidos comum a esse domínio. É provável que exista um número limitado, embora grande, de domínios de proteínas, que se misturaram e combinaram durante a evolução para gerar a diversidade de proteínas observadas nos organismos atualmente.

Já foram caracterizados muitos domínios de proteínas e suas funções moleculares foram determinadas. A sequência de um gene recém-identificado pode ser rastreada em comparação com um banco de dados de domínios conhecidos. Se a sequência de gene codifica um ou mais domínios cujas funções foram previamente determinadas, a função do domínio pode fornecer informações importantes sobre uma possível função do novo gene.

Conceitos

A função de um gene desconhecido pode ser determinada em alguns casos pelo achado de genes com uma sequência semelhante cuja função é conhecida. A função de um gene também pode ser determinada por meio da identificação de domínios funcionais na proteína que ele codifica.

Checagem dos conceitos 5

Qual a diferença entre ortólogos e parálogos?

a. Os ortólogos são sequências homólogas, os parálogos são sequências análogas.
b. Os ortólogos são mais semelhantes que os parálogos.
c. Os ortólogos estão no mesmo organismo, enquanto os parálogos estão em organismos diferentes.
d. Os ortólogos estão em organismos diferentes, os parálogos estão no mesmo organismo.

Expressão gênica e microarranjos

Muitos indícios importantes sobre a função do gene provêm do conhecimento de quando e onde os genes são expressos. O desenvolvimento de microarranjos permitiu que a expressão de milhares de genes seja monitorada simultaneamente.

Os **microarranjos** dependem da hibridização de ácido nucleico (ver *Capítulo 19*), na qual um fragmento de DNA é usado como uma sonda para encontrar sequências complementares (**Figura 20.9**). Vários fragmentos de DNA conhecidos são fixados em um suporte sólido em um padrão de organização, ou arranjo, em geral, uma série de pontos. Estes fragmentos de DNA (as sondas) correspondem aos genes conhecidos de um organismo específico.

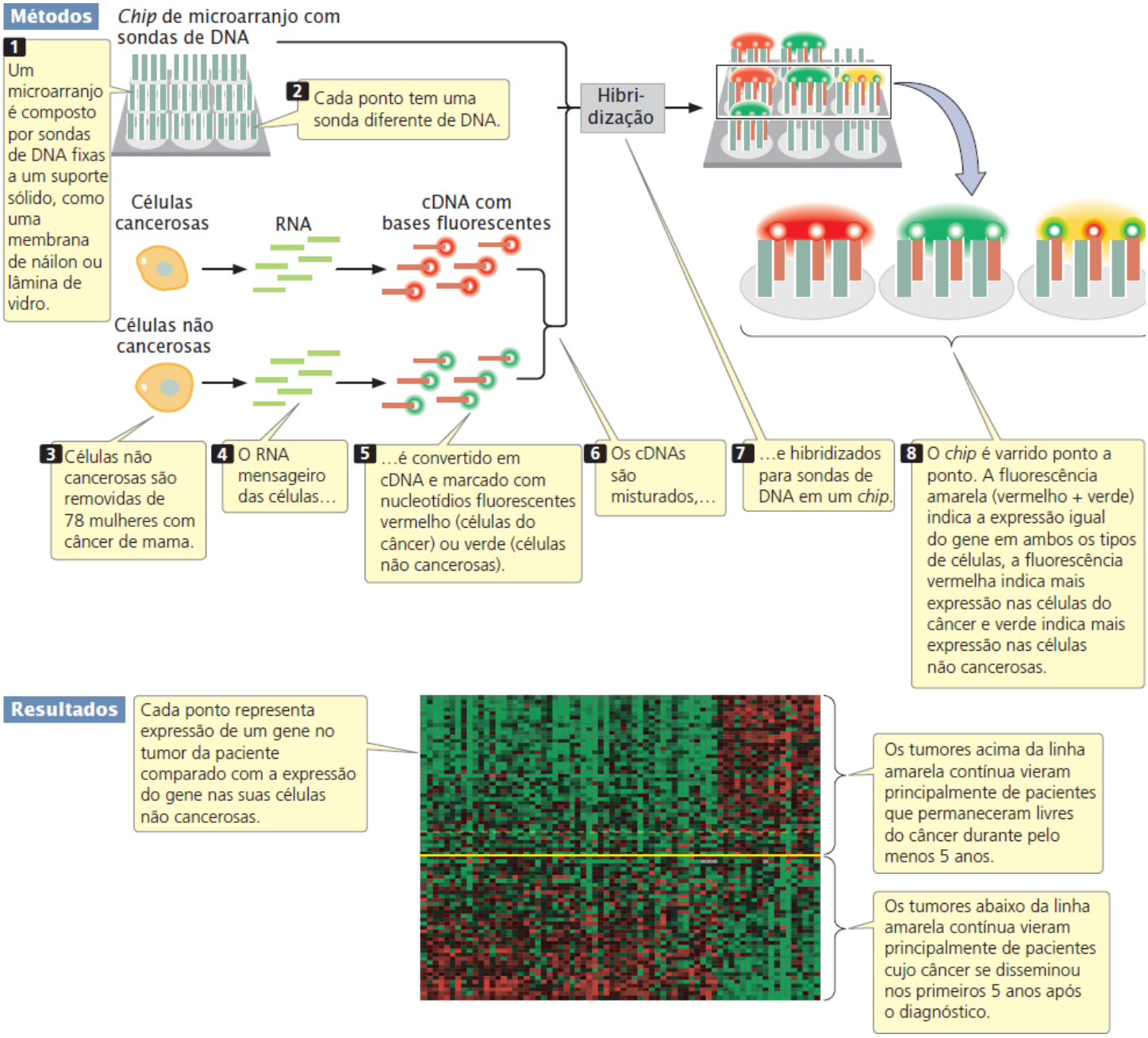
Após o microarranjo ter sido construído, o mRNA, DNA ou cDNA isolado das células experimentais é marcado com nucleotídios fluorescentes e aplicados ao arranjo. Qualquer uma das moléculas de DNA ou RNA complementares às sondas no arranjo vão hibridizar com elas e emitir fluorescência, que pode ser detectada por um escâner automático. Qualquer arranjo contendo dezenas de milhares de sondas pode ser aplicado a uma lâmina de vidro ou *wafer* de silicone de apenas alguns centímetros quadrados de tamanho.

Uso dos microarranjos. Usados com cDNA, os microarranjos podem fornecer informações sobre a expressão de milhares de genes, permitindo que os cientistas estudem quais genes estão ativos em tecidos específicos. Eles também podem ser usados para investigar como a expressão gênica muda no curso dos processos biológicos como o desenvolvimento ou progresso de doenças. Por exemplo, o câncer de mama afeta 1 de 10 mulheres nos EUA, e metade destas morrem. O tratamento atual depende de vários fatores, incluindo a idade da mulher, o tamanho do tumor, as características das células tumorais e se o câncer já se disseminou para os linfonodos próximos. Muitas mulheres cujo câncer não se espalhou são tratadas com a remoção do tumor e irradiação, mesmo assim, o câncer tardio reaparece em algumas das mulheres tratadas. Essas mulheres podem se beneficiar com um tratamento mais agressivo quando o câncer é detectado pela primeira vez.

Usando os microarranjos, os pesquisadores examinaram os padrões de expressão de 25.000 genes de tumores primários de 78 mulheres jovens com câncer de mama (ver Figura 20.9). O RNA mensageiro das células cancerosas e não cancerosas foi convertido em cDNA e marcado com nucleotídios vermelhos fluorescentes e verdes fluorescentes, respectivamente. Os cDNAs marcados foram misturados e hibridizados em um *chip* de DNA, que tinha sondas de DNA de diferentes genes. A hibridização dos cDNA vermelhos (câncer) e verdes (não câncer) foi proporcional às quantidades relativas de mRNA nas amostras. A fluorescência de cada ponto foi avaliada com varredura microscópica e aparecia como uma única cor. O vermelho indica a expressão aumentada de um gene nas células cancerosas em relação ao mesmo nas células não cancerosas (mais cDNA marcado com vermelho hibridiza), enquanto verde indica a expressão reduzida de um gene nas células cancerosas em relação ao mesmo nas células não cancerosas (mais cDNA marcado com verde hibridiza). O amarelo indica a expressão igual em ambos os tipos de células (hibridização igual dos cDNA marcados com verde e vermelho) e a ausência de cor indica a não expressão nos dois tipos de células.

Experimento

Pergunta: A variação na expressão gênica, detectada por microarranjos pode ser usada para prever a recorrência de câncer de mama?



Conclusão: Foram identificados setenta genes cujos padrões de expressão previam com exatidão a recorrência de câncer de mama nos primeiros 5 anos após o tratamento.

Figura 20.9 Os microarranjos podem ser usados para examinar a expressão gênica associada à progressão da doença. (Reimpressa com permissão de Macmillan Publishers Ltd: van'T Veer, Laura J et al., "Gene expression profiling predicts clinical outcome of breast cancer," NATURE 405:532, copyright 2002.)

Em 34 das 78 pacientes, o câncer tardio se disseminou para outros sítios; as outras 44 pacientes permaneceram sem câncer de mama por 5 anos após seu diagnóstico inicial. Os pesquisadores identificaram um subconjunto de 70 genes para os quais os padrões de expressão nos tumores iniciais previram com exatidão se o câncer se disseminaria posteriormente (ver Figura 20.9). Este grau de previsão foi muito maior que os das medidas tradicionais, que se baseavam no tamanho e na histologia do tumor.

Os pesquisadores também usaram microarranjos para examinar a expressão dos microRNAs (miRNAs) nos cânceres humanos. Pesquisa recente indica que miRNAs são expressos com frequência anormal no tecido canceroso e podem contribuir para a progressão do câncer (ver *Capítulo 23*). Por exemplo, um estudo usando microarranjos descobriu que vários miRNAs tinham a expressão aumentada no tecido do câncer do colo uterino, comparados com o tecido normal do colo, enquanto outros miRNAs tinham a expressão reduzida (**Figura 20.10**). Outros estudos usando microarranjos demonstraram que a expressão do miRNA está associada à resistência dos tumores a quimioterapia e radioterapia e que a expressão do miRNA pode ser usada para prever a resposta de alguns tumores ao tratamento anticâncer. Os resultados destes estudos sugerem que os dados de expressão gênica obtidos a partir dos microarranjos podem ser uma ferramenta poderosa para determinar a natureza da pesquisa e tratamento do câncer. Os produtos dos genes que mostram diferenças na expressão estão sendo examinados como possíveis

alvos para a terapia medicamentosa. Resolva o Problema 34

Foram criados microarranjos que permitem a detecção de alelos específicos, SNPs e até proteínas específicas. O mais importante, nem todas as moléculas de DNA se ligam da mesma forma aos microarranjos e, às vezes, os microarranjos superestimam ou subestimam a expressão de genes específicos. Assim, é desejável verificar os resultados dos microarranjos com outros métodos.

Conceitos

Os microarranjos, compostos de sondas de DNA conectadas a um suporte fixo, podem ser usados para determinar quais sequências de RNA e DNA estão presentes em uma mistura de ácidos nucleicos. Eles são capazes de determinar quais moléculas de RNA estão sendo sintetizadas e podem ser usados para examinar mudanças na expressão gênica.

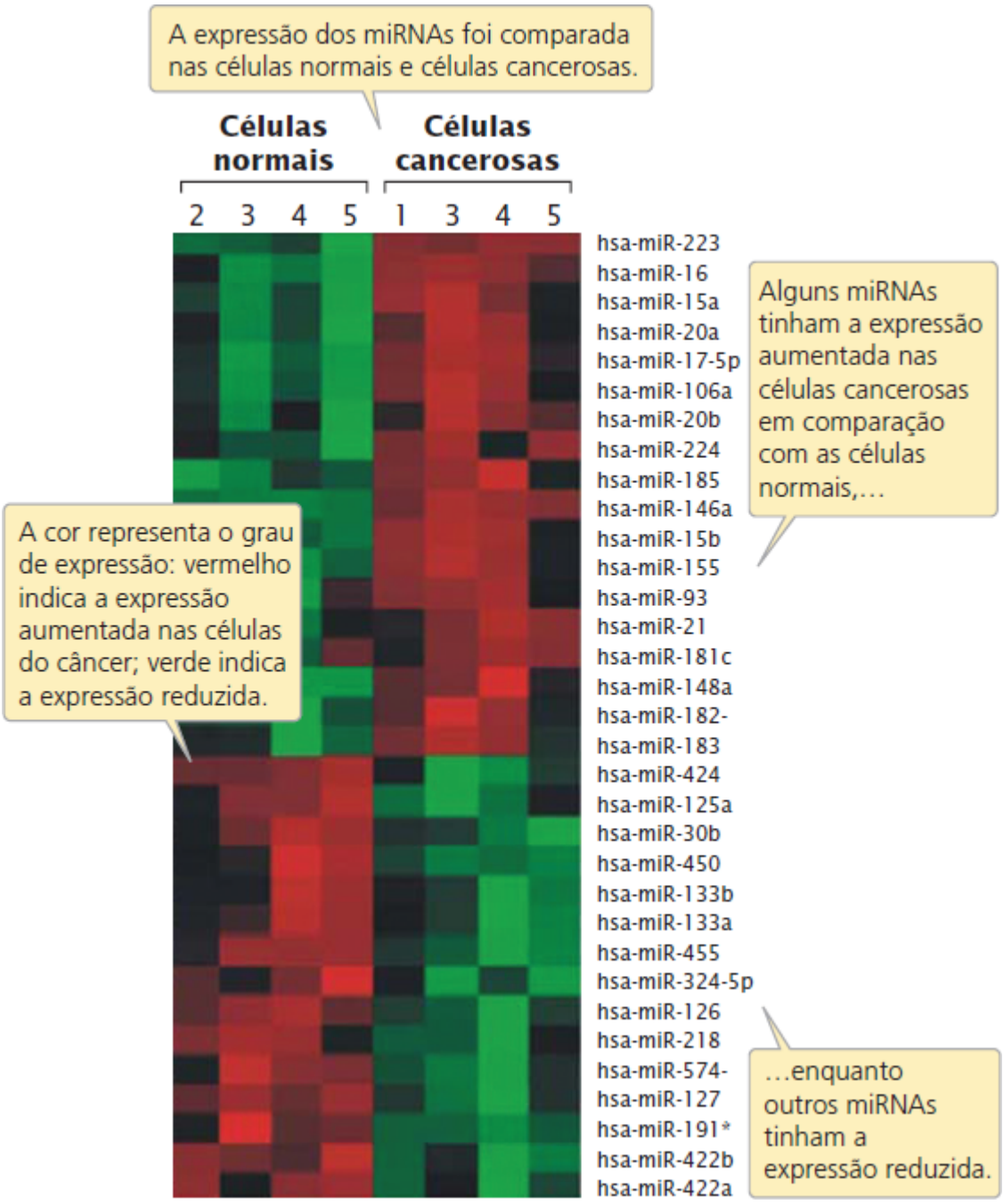


Figura 20.10 Os microarranjos são usados para comparar a expressão de miRNAs nas células do câncer do colo uterino comparadas com as células normais do colo uterino. (Adaptada de X. Wang, S. Tang, S-Y. Le, R. Lu, J.S. Rader et al. [2008] Aberrant Expression of Oncogenic and Tumor-Suppressive Micro RNAs in Cervical Cancer Is Required for Cancer Cell Growth. PLoSONE 3[7]: e2557. doi:10.1371/journal).

Expressão gênica e sequências repórter

Os padrões de expressão gênica também podem ser determinados visualmente ao usar uma sequência repórter. Nesta abordagem, os fragmentos genômicos são primeiro clonados em BACs ou outros vetores capazes de segurar a região codificadora de um gene mais suas sequências reguladoras. A região codificadora de um gene cuja expressão deve ser estudada é substituída por uma sequência repórter, que codifica um produto de fácil observação. Por exemplo, uma sequência repórter muito usada codifica uma proteína verde fluorescente (GFP) na água-viva. O BAC é então inserido em um embrião, criando um organismo transgênico. A sequência reguladora do gene clonado garante que ele é expresso no momento adequado e no tecido adequado dentro do organismo transgênico. O produto da expressão gênica é um pigmento verde fluorescente, facilmente observado (**Figura 20.11**).

Essa técnica está sendo usada para estudar os padrões de expressão dos genes que afetam a função do cérebro. No projeto Gene Expression Nervous System Atlas (GENSAT), os cientistas estão substituindo sistematicamente as regiões codificadoras de centenas de genes com a sequência repórter GFP e observando seus padrões de expressão nos camundongos transgênicos. O objetivo é produzir um atlas abrangente da expressão gênica no cérebro do camundongo. Este projeto já lançou luz sobre onde são expressos vários genes importantes transtornos neurológicos hereditários. Outro projeto de genômica funcional, o Allen Brain Atlas, compilou informações sobre os padrões de expressão de 20.000 genes no cérebro do camundongo.

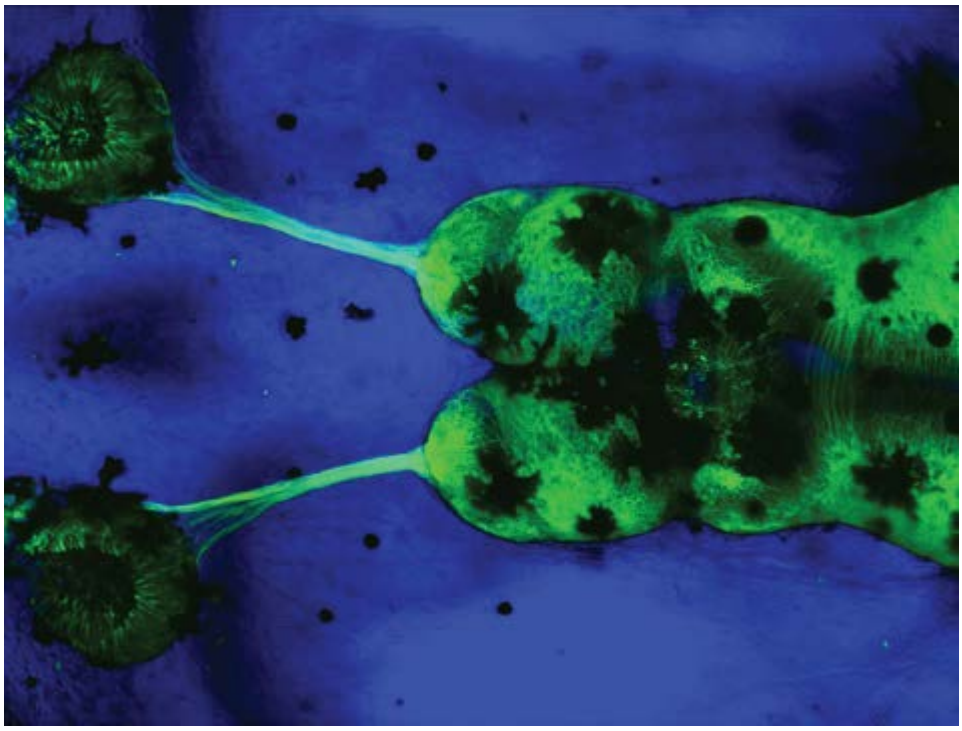


Figura 20.11 Uma sequência repórter pode ser usada para examinar a expressão de um gene. A expressão do gene da β-tubulina neural específica no cérebro de um girino é revelada pelo pigmento verde fluorescente. (Foto de Miranda Gomperts, cortesia Enrique Amaya, The Amaya Lab.)

Mutagênese genômica global

Como foi discutido no Capítulo 18, um dos melhores métodos para determinar a função de um gene é examinar os fenótipos de organismos individuais com uma mutação no gene. Tradicionalmente, foram mapeados os genes que codificam variações de ocorrência natural em um fenótipo, os genes causadores foram isolados e seus produtos foram estudados. Mas este procedimento foi limitado pelo número de mutações que ocorrem naturalmente pela dificuldade em mapear os genes com um número limitado de marcadores cromossômicos. O número destas mutações pode ser aumentado pela exposição a agentes mutagênicos, e a exatidão do mapeamento aumenta muito com a disponibilidade de marcadores moleculares mapeados, como os RFLPs, microssatélites, STSs, ESTs e SNPs. Estes dois métodos, indução aleatória de mutações com base no genoma inteiro e o mapeamento com marcadores moleculares, estão acoplados e automatizados em uma **triagem de mutagênese**.

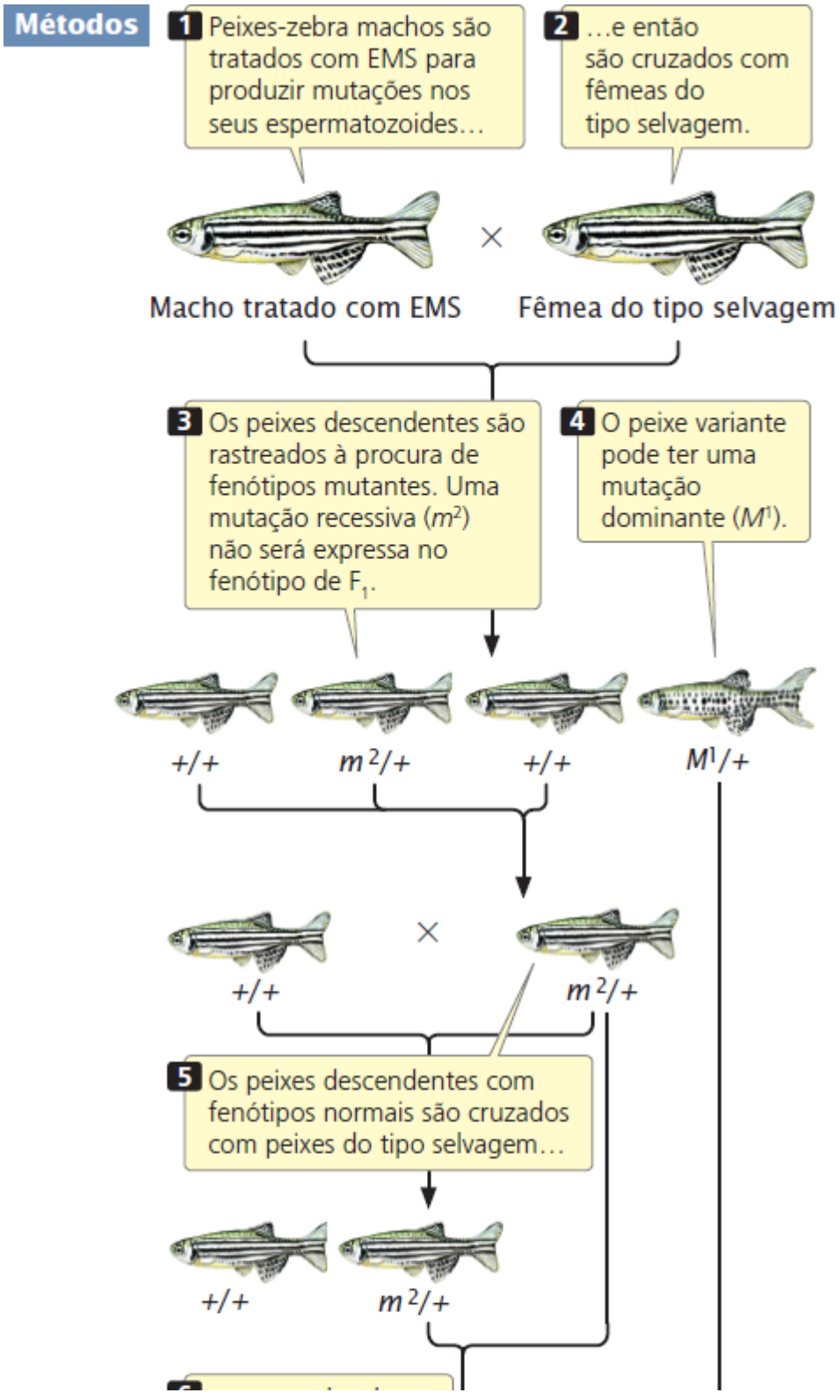
As triagens de mutagênese do genoma global podem ser usadas para todos os genes que afetam uma função ou traço específico. Por exemplo, as triagens de mutagênese dos camundongos estão sendo usadas para identificar os genes com papéis na função cardiovascular. Quando estes genes estão localizados nos camundongos, são feitas pesquisas de homologia para determinar se existem genes semelhantes nos seres humanos. Estes genes podem ser estudados para melhor compreender a doença cardíaca nos humanos.

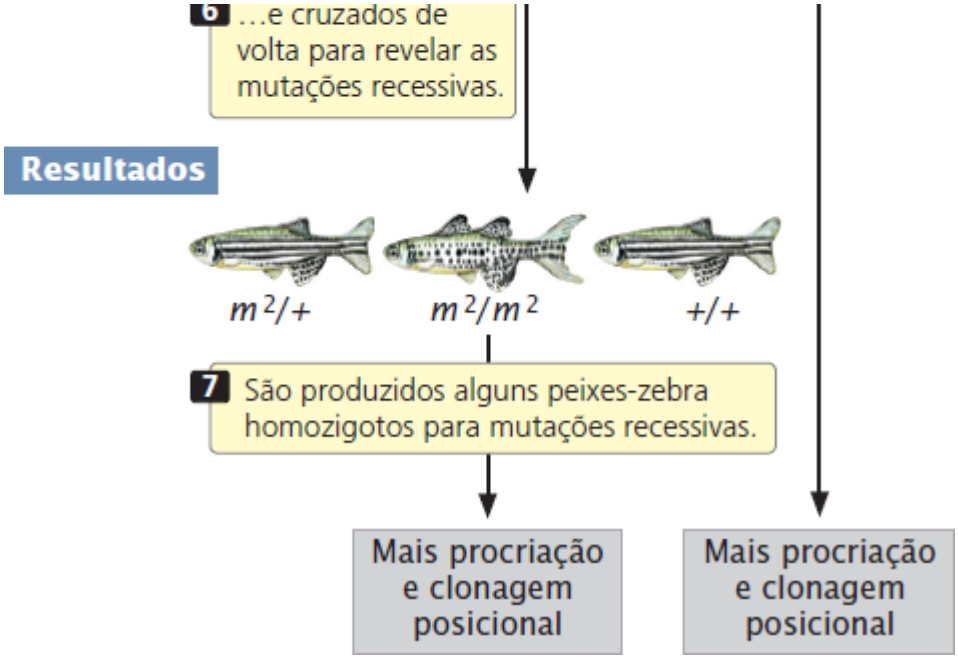
Para fazer uma triagem de mutagênese, são induzidas mutações aleatórias em uma população de organismos, criando nos fenótipos. As mutações são induzidas ao expor os organismos a radiação, um mutágeno químico (ver *Capítulo 18*) ou elementos

de transposição (sequências de DNA inseridas aleatoriamente no DNA, ver *Capítulo 18*). A **Figura 20.12** ilustra um procedimento para uma típica triagem de mutagênese. Aqui, os peixes-zebra machos são tratados com etilmetilsulfonato (EMS), uma substância química que induz mutações na linhagem celular (nos espermatozoides). Os machos tratados são cruzados com peixes-fêmea do tipo selvagem. Alguns dos descendentes serão heterozigotos para as mutações induzidas pelo EMS; os descendentes serão examinados à procura de fenótipo mutante que possa ser o produto das mutações dominantes expressas neste peixe heterozigoto. As mutações recessivas não serão expressas nos descendentes F_1, mas podem ser reveladas com cruzamentos adicionais.

Experimento

Pergunta: Como os genes que codificam um traço ou função específica podem ser identificados?





Conclusão: A triagem de mutagênese produz peixe sem uma mutação que afeta o traço. Estes peixes são analisados para identificar a mutação.

Figura 20.12 Os genes que afetam uma característica ou função específica podem ser identificados pela triagem de mutagênese de genoma global. Nesta ilustração, M^1 representa uma mutação dominante e m^2 representa uma mutação recessiva.

O peixe com fenótipos variantes sofre outros experimentos de procriação para verificar se cada fenótipo variante é, de fato, causado por uma mutação de gene único na qual a mutação que ocorre é desconhecida. Entretanto, o gene pode ser localizado por clonagem posicional (ver *Capítulo 19*).

As triagens de mutagênese são usadas para estudar os genes que controlam o desenvolvimento dos vertebrados. Uma equipe de geneticistas de desenvolvimento produziu milhares de mutações no peixe-zebra que afetam o desenvolvimento e estão localizadas sistematicamente e caracterizadas nos *loci* onde as mutações ocorrem. Os peixes-zebra são modelos genéticos ideais para este tipo de estudo porque se reproduzem rapidamente, são cultivados com facilidade no laboratório e têm embriões transparentes nos quais é fácil localizar as deformidades de desenvolvimento. Esta pesquisa já identificou vários genes importantes no desenvolvimento embrionário, muitos dos quais têm correspondentes nos humanos.

Conceitos

A triagem de mutagênese de genoma global acoplado com a clonagem posicional pode ser usada para identificar os genes que afetam uma característica ou função específica.

Checagem dos conceitos 6

Qual é a ordem correta das etapas em uma triagem de mutagênese?

a. Clonagem posicional, mutagênese, identificação dos mutantes, verificação da base genética.
b. Mutagênese, clonagem posicional, identificação dos mutantes, verificação da base genética.
c. Mutagênese, identificação dos mutantes, verificação da base genética, clonagem posicional.
d. Identificação dos mutantes, clonagem posicional, mutagênese, verificação da base genética.

20.3 A genômica comparativa estuda como os genomas evoluem

Os projetos de sequenciamento de genoma fornecem informações detalhadas sobre o conteúdo e a organização dos genes em diferentes espécies e até em diferentes membros da mesma espécie, permitindo inferências sobre como os genes funcionam e como os genomas evoluem. Eles também fornecem informações importantes sobre as relações evolutivas entre os organismos e sobre fatores que influenciam a velocidade e a direção da evolução. A **genômica comparativa** é o campo da genômica que compara as semelhanças e diferenças no conteúdo, função do gene e a organização entre os genomas de diferentes organismos.

Genomas procarióticos

Já foram sequenciados mais de 2.000 genomas procarióticos. A maioria destes genomas é composta por um único cromossomo circular. Entretanto, existem exceções, como o *Vibrio cholerae*, a bactéria que provoca cólera, que tem dois cromossomos circulares e *Borrelia burgdorferi*, que tem um grande cromossomo linear e 21 cromossomos menores.

Tamanho do genoma e número de genes. A quantidade total de DNA nos genomas procarióticos varia de 490.885 pb no *Nanoarchaeum equitans*, um archaea que vive inteiramente dentro de outra archaea, a 9.105.828 pb na *Bradyrhizobium japonicum*, uma bactéria do solo (**Quadro 20.1**). Embora esta faixa no tamanho do genoma pareça ser grande, é muito menor que a imensa faixa de tamanhos de genomas observados nos eucariotos, que podem variar de alguns milhões de pares de bases a centenas de bilhões de pares de base. *Escherichia coli*, a bactéria mais usada para estudos genéticos, tem um tamanho de genoma bem típico de 4,6 milhões de pares de bases. Archaea e bactérias são semelhantes em suas faixas de tamanho do genoma. De forma surpreendente, o tamanho do genoma também varia muito em algumas espécies, por exemplo, diferentes cepas de *E. coli* apresentam tamanhos diferentes do genoma em mais de 1 milhão de pares de bases.

Entre os procariotos, o número de genes varia em geral de 1.000 a 2.000, mas algumas espécies têm mais de 6.700 e outras apenas 480. O mais interessante, a densidade dos genes é constante em todas as espécies, com uma média de um gene a cada 1.000 pb. Assim, os procariotos com genomas maiores terão mais genes, ao contrário dos eucariotos, para os quais existe pouca associação entre o tamanho do genoma e o número de genes (ver a próxima seção). Os fatores evolutivos que determinam o tamanho dos genomas procarióticos (assim como os genomas eucarióticos) ainda não são conhecidos. Entretanto, o tempo necessário para a divisão celular é maior nos organismos com genomas maiores por causa do tempo necessário para replicar o DNA antes da divisão. Assim, a seleção pode favorecer genomas maiores nos organismos que ocupam ambientes onde a reprodução rápida é vantajosa.

Quadro 20.1 Características dos genomas procarióticos representativos que foram completamente sequenciados.

Espécie	Tamanho (milhões de pares de bases)	Número de genes previstos
Archaea		
Archaeoglobus fulgidus	2,18	2.407
Methanobacterium thermoautotrophicum	1,75	1.869
Nanoarchaeum equitans	0,490	536
Eubactérias		
Bacillus subtilis	4,21	4.100
Bradyrhizobium japonicum	9,11	8.317
Escherichia coli	4,64	4.289
Haemophilus influenzae	1,83	1.709
Mycobacterium tuberculosis	4,41	3.918
Mycoplasma genitalium	0,58	480
Staphylococcus aureus	2,88	2.697
Vibrio cholerae	4,03	3.828

Fonte: Dados do Genome Atlas of the Center for Biological Sequence Analysis, www.cbs.dtu.dk/services/GenomeAtlas.

Os procariotos com os menores genomas tendem a ser espécies que ocupam *habitats* restritos, como as bactérias que vivem dentro de outros organismos. O ambiente constante e as funções metabólicas fornecidas pelo organismo hospedeiro podem permitir que estas bactérias sobrevivam com menos genes. As bactérias com genomas maiores tendem a ocupar ambientes altamente complexos e variáveis, como o solo ou os nódulos de raízes das plantas. Nestes ambientes, os genes necessários

apenas em algumas situações podem ser úteis. Em ambientes mais complexos, onde os recursos são abundantes, pode ocorrer pouca necessidade de divisão rápida (o que favorece os genomas menores).

Um exemplo de bactéria de genoma grande que ocupa um ambiente complexo é a *Streptomyces coeliocolor*, uma bactéria filamentosa com um genoma de 8.677.507 pb. Esta bactéria é chamada de bactéria "escoteira" porque tem um conjunto variado de genes e, portanto, segue o lema dos escoteiros: sempre alerta. Além dos genes comuns de manutenção para as funções genéticas como replicação, transcrição e tradução, seu genoma tem vários genes para a degradação de carboidratos complexos, permitindo que ela consuma matéria em decomposição a partir de plantas, animais, insetos, fungos e outras bactérias. Tem genes para várias proteínas que produzem metabólitos secundários (produtos de degradação) que podem atuar como antibióticos e proteção contra dessecação e baixas temperaturas. Nesta espécie, um grande genoma com muitos genes parece ser benéfico no ambiente complexo que habita.

Transferência horizontal de genes. As bactérias têm vários mecanismos por meio dos quais elas podem ganhar ou perder DNA. O DNA pode ser perdido pela simples deleção e pode ser obtido pela duplicação gênica ou pela inserção de elementos genéticos de transposição. Outro mecanismo para obter informações genéticas novas é a transferência horizontal de genes, um processo pelo qual espécies bacterianas próximas ou não trocam informação genética periodicamente ao longo da evolução. Tal troca pode ocorrer por meio da captação da bactéria do DNA do ambiente (transformação) pela troca de plasmídios ou por vetores virais (ver *Capítulo 9*). A transferência horizontal de genes foi identificada há algum tempo, mas as análises de muitos genomas microbianos agora indicam que é mais abrangente do que se pensava antes. A ocorrência disseminada da transferência horizontal fez alguns biológicos se perguntarem se até existem espécies distintas nas bactérias (ver *Capítulo 9* para uma discussão mais detalhada sobre este assunto).

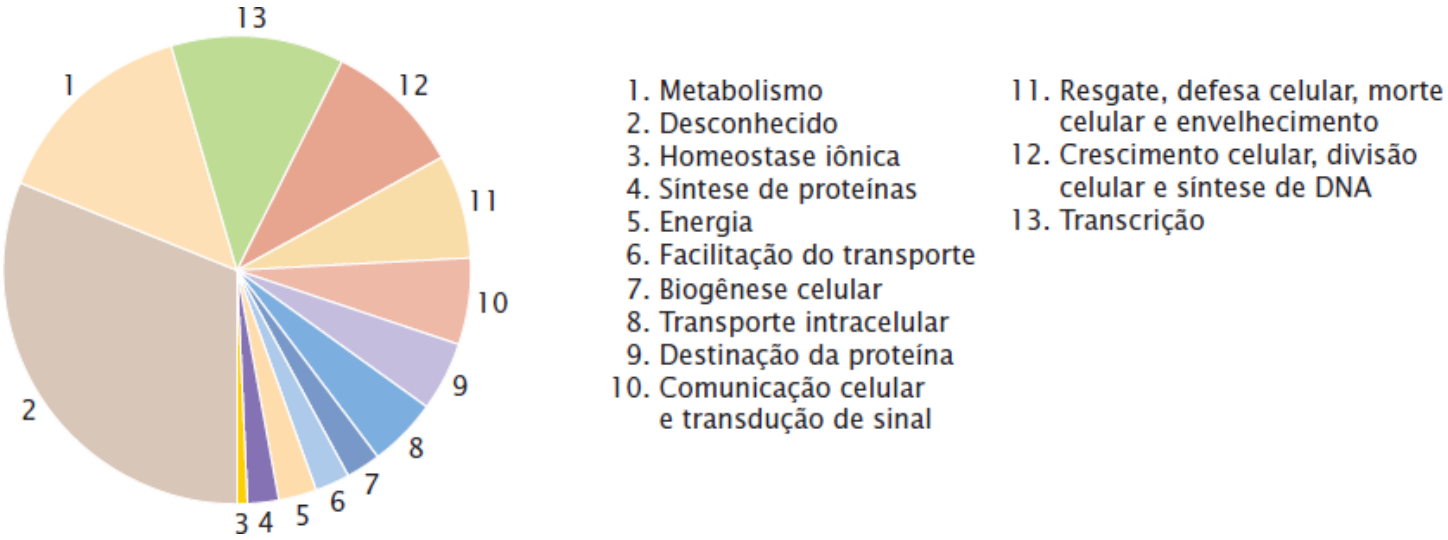


Figura 20.13 As funções de muitos genes nos procariotos não podem ser determinadas por comparação com genes em outros procariotos. A proporção do círculo ocupado por cada cor representa a proporção de genes que afetam várias funções conhecidas e desconhecidas na *E. coli*.

Função dos genes. Apenas metade dos genes identificados nos genomas procarióticos tem uma função específica. Quase um quarto dos genes não tem semelhança significativa da sequência a nenhum dos genes conhecidos em outras bactérias. O número de genes que codificam funções biológicas como transcrição e tradução tende a ser semelhante entre as espécies, mesmo quando seus genomas têm tamanhos muito diferentes. Esta semelhança sugere que estas funções são codificadas por um conjunto básico de genes que não varia entre as espécies. Ao contrário, o número de genes que participam de biossíntese, metabolismo de energia, transporte e funções reguladoras varia muito entre as espécies e tende a ser maior nas espécies com genomas maiores. As funções dos genes previstos (ou seja, genes identificados por programas de computador) e genes conhecidos na *E. coli* estão apresentadas na **Figura 20.13**. Uma parte substancial do DNA "extra" encontrada nos genomas bacterianos maiores é composta de genes parálogos que surgiram pela duplicação.

Conceitos

A genômica comparativa compara o conteúdo e a organização de sequências genômicas inteiras de diferentes organismos. Os genomas procarióticos são pequenos, em geral variando de 1 a 3 milhões de pares de bases de DNA, com milhares de genes. Os procariotos com os genomas menores tendem a ocupar *habitats* restritos, enquanto os procariotos com genomas maiores são encontrados em ambientes complexos.

A transferência horizontal de genes é importante para a evolução do genoma bacteriano.

Checagem dos conceitos 7

Qual é a relação entre o tamanho do genoma e o número de genes nos procariotos?

Genomas eucarióticos

Os genomas de mais de 150 organismos eucarióticos já foram completamente sequenciados, incluindo vários fungos e protistas, vários insetos, quase 30 espécies de plantas e numerosos vertebrados. Os eucariotos sequenciados incluem mamões, milho, arroz, sorgo, videira, bichos-da-seda, várias moscas-da-fruta, afídeos, mosquitos, anêmonas, camundongos, ratos, cães, vacas, cavalos, orangotangos, chimpanzés e humanos. Mesmo os genomas de alguns organismos extintos foram sequenciados, incluindo o mamute-lanoso e os neandertais. Centenas de genomas eucarióticos adicionais estão no processo de sequenciamento. É importante observar que, mesmo que os genomas destes organismos sejam "completamente sequenciados", muitas das sequências finais montadas têm lacunas e regiões de heterocromatina podem não ter sido sequenciadas. Assim, os tamanhos dos genomas eucarióticos são estimativas na maioria das vezes e o número de pares de bases para o tamanho do genoma de uma espécie em questão pode variar. É difícil prever o número de genes presentes em um genoma e ele pode variar, dependendo das considerações feitas e do programa usado para encontrar o gene.

Tamanho do genoma e número de genes. Os genomas dos organismos eucarióticos (**Quadro 20.2**) são maiores que os genomas dos procariotos e, em geral, os eucariotos multicelulares têm mais DNA do que os eucariotos simples, unicelulares como a levedura (ver *Capítulo 11*). Entretanto, não existe relação próxima entre o tamanho do genoma e a complexidade entre os eucariotos multicelulares. Por exemplo, o mosquito (*Anopheles gambiae*) e a mosca-da-fruta (*Drosophila melanogaster*) são insetos com complexidade estrutural semelhante, mas o mosquito tem 60% a mais de DNA do que a mosca-da-fruta.

Em geral, os genomas eucarióticos também têm mais genes que os dos procariotos (embora algumas bactérias maiores tenham mais genes que as leveduras unicelulares) e os genomas dos eucariotos multicelulares têm mais genes que os genomas dos eucariotos unicelulares. Ao contrário das bactérias, não existe correlação entre o tamanho do genoma e o número de genes nos eucariotos. O *número* de genes entre os eucariotos multicelulares também não está claramente relacionado com a complexidade fenotípica: os humanos têm mais genes que os invertebrados, mas apenas duas vezes a mais que as moscas-da-fruta e pouco menos que a planta *A. thaliana*. O nematódeo *C. elegans* tem mais genes que a *D. melanogaster*, mas é menos complexo. Além disso, o peixe-bola ou baiacu tem apenas um décimo da quantidade de DNA existente nos humanos e os camundongos, mas próximos em quantidade de genes. Os genomas eucarióticos têm múltiplas cópias de muitos genes, indicando que a duplicação de genes é um processo importante na evolução do genoma.

A maioria do DNA nos organismos multicelulares não é codificada e muitos genes são interrompidos por íntrons. Nos eucariotos mais complexos, o número e o comprimento dos íntrons são maiores.

Duplicações segmentares e famílias de multigenes. Muitos genomas eucarióticos, em especial dos organismos multicelulares, são preenchidos com **duplicações segmentares**, regiões com mais de 1.000 pb quase idênticas na sequência. Por exemplo, cerca de 4% do genoma humano têm duplicações segmentares. Na maioria das duplicações segmentares, as duas cópias são encontradas no mesmo cromossomo (uma duplicação intracromossômica), mas em outros, as duas cópias são encontradas em diferentes cromossomos (uma duplicação intercromossômica). No genoma humano, o tamanho médio das duplicações segmentares é de 15.000 pb.

As duplicações segmentares surgem a partir de processos que geram duplicações cromossômicas, como um *crossing over* desigual (ver *Capítulo 8*). Após surgir uma duplicação segmentar, ela promove mais duplicação pela ocorrência do desalinhamento entre as regiões duplicadas. Esta duplicação é importante na evolução ao dar origem a novos genes. Após surgir uma duplicação segmentar, a cópia original do gene pode continuar sua função enquanto a nova cópia sofre mutação. Estas mudanças podem consequentemente levar a uma nova função. A importância da duplicação do gene na evolução do genoma é demonstrada pelo grande número de famílias de multigenes que existem em muitos genomas eucarióticos. Uma **família de multigenes** é um grupo de genes evolutivamente próximos que surgem por meio de evolução repetida de um gene ancestral. Por exemplo, a família do gene da globina nos humanos é composta por 13 genes que codificam moléculas semelhantes à globina, a maioria dos quais produz proteínas que transportam oxigênio. Até um exemplo mais espetacular é a família de multigenes olfatórios humanos, composta por cerca de 1.000 genes que codificam as moléculas receptoras olfatórias usadas no nosso olfato.

Quadro 20.2 Características dos genomas eucarióticos representativos que foram completamente sequenciados.

Espécie	Tamanho do genoma (milhões de pares de bases)	Número de genes previstos
Saccharomyces cerevisiae (levedura)	12	6.144
Physcomitrella patens (musgo)	480	38.354
Arabidopsis thaliana (planta mostarda)	125	25.706
Zea mays (milho)	2.400	32.000
Hordeum vulgare (cevada)	5.100	26.159
Caenorhabditis elegans (nematódeo)	103	20.598
Drosophila melanogaster (mosca-da-fruta)	170	13.525
Anopheles gambiae (mosquito)	278	14.707
Danio rerio (peixe-zebra)	1.465	22.409
Takifugu rubripes (baiacu)	329	22.089
Xenopus tropicalis (rã)	1.510	18.429
Anolis carolinesis (lagarto)	1.780	17.792
Mus musculus (camundongo doméstico)	2.627	26.762
Pan troglodytes (chimpanzé)	2.733	22.524
Homo sapiens (humano)	3.223	Cerca de 20.000

Fonte: Ensembl Web site: http://useast.ensembl.org/index.html e plants.ensembl.org/index.html

DNA não codificador. A maioria dos organismos eucarióticos tem grande quantidade de DNA que não codifica proteínas. Por exemplo, apenas 1,5% do genoma humano é composto de DNA que especifica diretamente os aminoácidos das proteínas. A função das sequências de DNA remanescentes, chamadas de DNA não codificador, tem sido questionada. Algumas pesquisas sugerem que parte deste DNA não tem função. Por exemplo, Marcelo Nobrega e seus colaboradores criaram camundongos geneticamente modificados que não tinham uma grande região cromossômica sem genes codificadores de proteínas (chamada de **deserto gênico**). Em um experimento eles criaram camundongos sem um deserto gênico de 1.500.000 pb do cromossomo 3 de camundongo; em outro, eles criaram camundongos sem um deserto gênico de 845.000 pb do cromossomo 19. Incrivelmente, estes camundongos eram saudáveis e indistinguíveis dos camundongos do controle. Os pesquisadores concluíram que grandes regiões do genoma dos mamíferos podem ser eliminadas sem grandes efeitos fenotípicos e podem, de fato, ser supérfluas.

Entretanto, outra pesquisa sugeriu que os desertos gênicos podem ter sequências com papel funcional. Por exemplo, estudos de associação genômica global demonstraram que as sequências de DNA contidas em um deserto gênico no cromossomo 9 humano estão associadas a coronariopatia e outros estudos demonstraram a existência de 33 acentuadores neste deserto gênico.

Em 2002, o projeto *Encyclopedia of DNA Elements* (ENCODE) foi criado para determinar se o DNA não codificado tinha alguma função. Os pesquisadores catalogaram todos os nucleotídios no genoma que fornecem alguma função, incluindo sequências codificadoras de proteínas e moléculas de RNA e os que servem como sítios de controle para expressão gênica. Este projeto de 10 anos foi feito por uma equipe de mais de 400 cientistas ao redor do mundo. Em uma série de artigos publicados em 2012, ENCODE concluiu que pelo menos 80% do genoma humano está envolvido em algum tipo de função gênica. Muitas das sequências funcionais eram compostas por sítios onde as proteínas se ligavam e influenciavam a expressão dos genes. Antes deste estudo, parte do genoma era considerada “DNA lixo” sem função, mas o estudo ENCODE modificou significativamente esta visão e sugere que existe pouco DNA não funcional no genoma humano.

Elementos de transposição. Uma parte substancial dos genomas da maioria dos organismos multicelulares é composta de

sequências moderadamente e muito repetitivas (ver *Capítulo 11*) e a porcentagem de sequências repetitivas é maior que nas espécies com genomas maiores (**Quadro 20.3**). Parece que a maioria destas sequências repetitivas surge por meio da transposição. No genoma humano, 45% do DNA é derivado de elementos de transposição, muitos dos quais são defeituosos e não são mais capazes de se mover. No milho, 85% do genoma são derivados dos elementos de transposição.

Quadro 20.3 Porcentagem de genoma composto por repetições intercaladas derivadas dos elementos de transposição.

Organismo	Porcentagem do genoma
Arabidopsis thaliana (planta)	10,5
Zea mays (milho)	85,0
Caenorhabditis elegans (nematódeo)	6,5
Drosophila melanogaster (mosca-da-fruta)	3,1
Takifugu rubripes (baiacu)	2,7
Homo sapiens (humano)	44,4

Diversidade das proteínas. Apesar de apenas um pequeno aumento no número de genes, os vertebrados têm uma diversidade de proteínas consideravelmente maior do que os invertebrados. Uma forma de medir a quantidade de diversidade de proteínas é contar o número de domínios de proteínas, que são partes características das proteínas associadas a uma função. Os genomas de vertebrados não codificam mais domínios de proteínas do que os genomas dos invertebrados; por exemplo, existem 1.262 domínios nos humanos em comparação aos 1.035 nas moscas-da-fruta (**Quadro 20.4**). Entretanto, os domínios existentes nos humanos são reunidos em mais combinações, resultando em muito mais tipos de proteínas.

Genes homólogos. Uma tendência clara e incrível observada nos genomas eucarióticos é o grau de homologia entre os genes encontrados até mesmo em espécies distantes. Por exemplo, os camundongos e os seres humanos têm cerca de 99% dos seus genes em comum. Cerca de 50% dos genes nas moscas-da-fruta são homólogos aos genes nos humanos, e até nas plantas, cerca de 18% dos genes são homólogos aos encontrados nos humanos.

Colinearidade entre genes correlatos. Uma das características da evolução do genoma revelada pela comparação das sequências de genes de diferentes organismos é que muitos genes estão presentes na mesma ordem nos genomas relacionados, um fenômeno chamado colinearidade. O motivo para a colinearidade entre os genomas é que eles descendem de um genoma ancestral comum e as forças evolutivas mantiveram a mesma ordem de genes nos genomas dos descendentes. Os estudos genômicos das gramíneas, as plantas da família Poaceae, ilustram o princípio da colinearidade.

Quadro 20.4 Número de domínios de proteínas estimados codificados por alguns genomas eucarióticos.

Espécie	Número de domínios de proteínas previstos
Saccharomyces cerevisiae (levedura)	851
Arabidopsis thaliana (planta)	1.012
Caenorhabditis elegans (nematódeo)	1.014
Drosophila melanogaster (mosca-da-fruta)	1.035
Homo sapiens (humano)	1.262

Fonte: Number of genes and protein-domain families from the International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome, Nature 409:860-921, Table 23, 2001.

As gramíneas incluem mais de 10.000 espécies, incluindo plantações economicamente importantes, como arroz, trigo, cevada, milho, painço, aveia e sorgo. Juntas, as gramíneas compõem cerca de 60% da produção mundial de alimentos. Os genomas destas espécies variam muito em tamanho e número de cromossomos. Por exemplo, o número de cromossomos nas

gramíneas varia de 4 a 26, o genoma do arroz é composto por apenas 460 milhões de pares de bases, enquanto o genoma do trigo tem 17 bilhões de pares de bases. Apesar dessas grandes diferenças no número dos cromossomos e no tamanho do genoma, a posição e a ordem de muitos genes nos genomas são conservados. Com a evolução, as regiões de DNA entre os genes (regiões intergênicas) aumentaram, reduziram e sofreram rearranjos, enquanto os próprios genes ficaram relativamente constantes em ordem e conteúdo. A **Figura 20.14** mostra um exemplo de relação colinear dos genes entre o arroz, o sorgo e *Brachypodium* (gramínea). Resolva o Problema 38

Conceitos

O tamanho do genoma varia muito entre as espécies eucarióticas. Para os organismos eucarióticos multicelulares, não existe uma relação clara entre a complexidade do organismo e a quantidade de DNA ou número de genes. Uma parte substancial do genoma nos organismos eucarióticos é composta por DNA repetitivo, parte deste é derivada dos elementos de transposição. Muitos genomas eucarióticos têm genes homólogos em comum e os genes estão na mesma ordem nos genomas dos organismos correlatos.

Checagem dos conceitos 8

As duplicações segmentares são importantes na evolução ao
a. Dar origem a novos genes e famílias de multigenes.
b. Manter o número de genes constante em um genoma.
c. Eliminar sequências repetitivas produzidas pela transposição.
d. Controlar o conteúdo de bases do genoma.

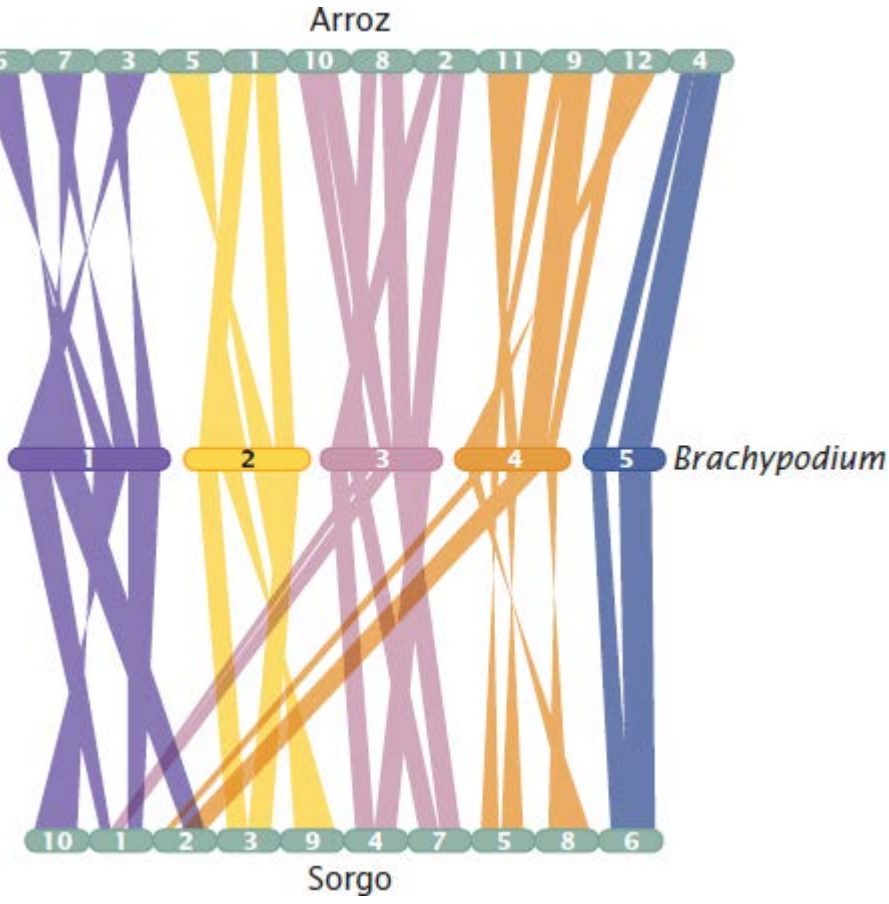


Figura 20.14 Relações colineares entre os blocos de genes encontrados no arroz, no sorgo e na *Brachypodium* (gramínea). Cada banda colorida representa um bloco de genes que é colinear entre os cromossomos das três espécies.

Genômica comparativa da *Drosophila*

A mosca-da-fruta *Drosophila melanogaster* é muito usada em genética. Seu genoma foi sequenciado em 2000, o segundo genoma animal a ser decifrado. Em 2007, os pesquisadores publicaram as sequências genômicas de 10 espécies adicionais de *Drosophila*. Combinado com os genomas já sequenciados de *D. melanogaster* e *D. pseudoobscura*, este esforço forneceu dados genômicos para 12 espécies de *Drosophila*, permitindo a análise detalhada da evolução deste grupo de insetos.

As 12 espécies de *Drosophila* sequenciadas são encontradas em todo o mundo e todas se originam de um ancestral comum há cerca de 60 milhões de anos (**Figura 20.15**). O tamanho do genoma no grupo varia de 130 milhões de pares de bases na *D. mojavensis* a 364 milhões de pares de bases da *D. virilis*. O número de genes que codificam proteínas varia menos, de 13.733 a 17.325. Muitos destes genes são encontrados em todas as 12 espécies; por exemplo, 77% dos 13.733 genes encontrados na *D. melanogaste*r têm homólogos em todas as outras espécies.

Os elementos de transposição são importantes na evolução nos genomas da *Drosophila*. A extensão de DNA que compõe o restante dos elementos de transposição varia de 2,7% do genoma na *D. simulans* a 25% do genoma da *D. ananassae*. A quantidade de DNA que é contribuição dos elementos de transposição é, em parte, responsável pela diferença no tamanho do genoma entre as espécies. Os rearranjos genômicos (inversões e translocações) foram frequentes na evolução deste grupo. A taxa de evolução varia em diferentes classes de genes: os genes que controlam o olfato, a imunidade e a resistência a inseticidas evoluíram mais rapidamente que outros genes.

Genoma humano

O genoma humano, razoavelmente típico dos genomas dos mamíferos, foi muito estudado e analisado por causa de sua importância para a saúde e a evolução humanas. Tem 3,2 bilhões de pares de base de comprimento, mas apenas 1,5% codificam proteínas. Os genes ativos estão separados por amplas regiões de DNA não codificador, parte do qual é composto por sequências repetidas derivadas dos elementos de transposição.

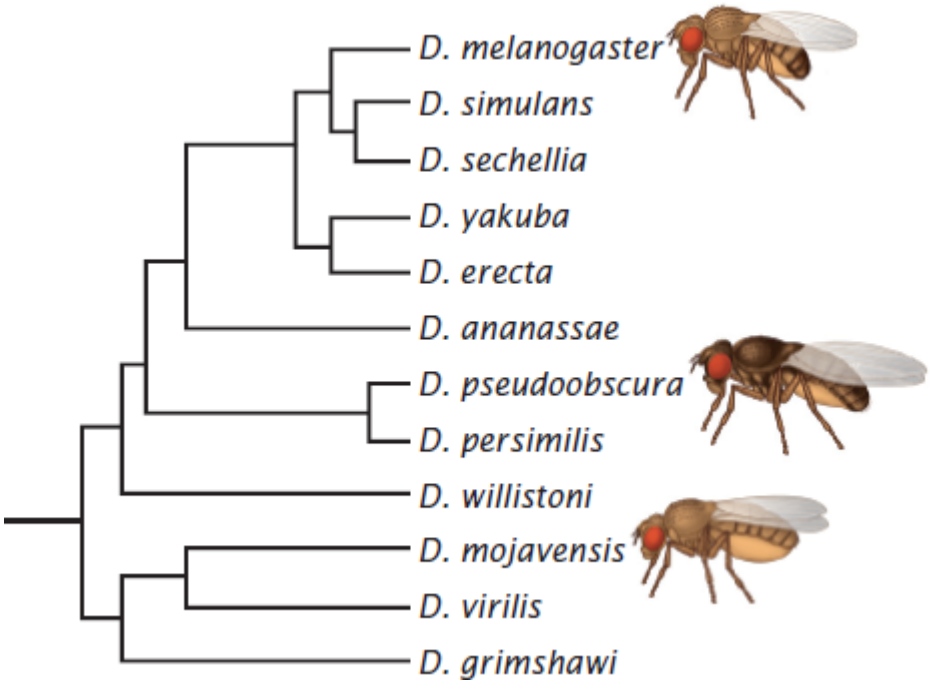


Figura 20.15 As doze espécies de *Drosophila* cujos genomas foram determinados e analisados divergiram a partir de um ancestral comum há aproximadamente 60 milhões de anos.

O gene médio no genoma humano tem aproximadamente 27.000 pb de comprimento, com cerca de 9 éxons (**Quadro 20.5**). (Um gene excepcional tem 234 éxons.) Os íntrons dos genes humanos são muito mais longos e existem mais deles do que em outros genomas (**Figura 20.16**). O genoma humano não codifica substancialmente mais domínios de proteínas (ver Quadro 20.4), mas os domínios são combinados de mais formas para produzir um proteoma relativamente diverso. As funções dos genes codificados pelo genoma humano são apresentadas na **Figura 20.17**. Um único gene codifica proteínas múltiplas por *splicing* ou recomposição alternativa; cada gene codifica, em média, dois ou três diferentes mRNAs.

A densidade gênica varia entre os cromossomos humanos, com os cromossomos 17, 19 e 22 apresentando as maiores densidades e os cromossomos X, Y, 4, 13 e 18 as menores densidades. Algumas proteínas codificadas pelo genoma humano que não são encontradas em outros animais incluem: as que afetam a função imunológica, o desenvolvimento neurológico, a estrutura e a função, as vias de sinalização intercelular e intracelular no desenvolvimento, hemostasia e apoptose.

Os elementos de transposição são muito mais comuns no genoma humano que no genoma do nematódeo, da planta e da mosca-da-fruta (ver Quadro 20.3). O genoma humano tem vários tipos de elementos de transposição, incluindo LINEs, SINEs, retrotranspósons e transpósons de DNA (ver *Capítulo 18*). Parece que a maioria é evolutivamente antiga e defeituosa, contendo mutações e deleções que os tornam incapazes de transposição.

Quadro 20.5 Características médias dos genes no genoma humano.

Característica	Média
Número de éxons	8,8
Tamanho do éxon interno	145 pb
Tamanho do íntron	3.365 pb
Tamanho da região 5′ não traduzida	300 pb
Tamanho da região 3′ não traduzida	770 pb
Tamanho da região codificadora	1.340 pb
Comprimento total do gene	27.000 pb

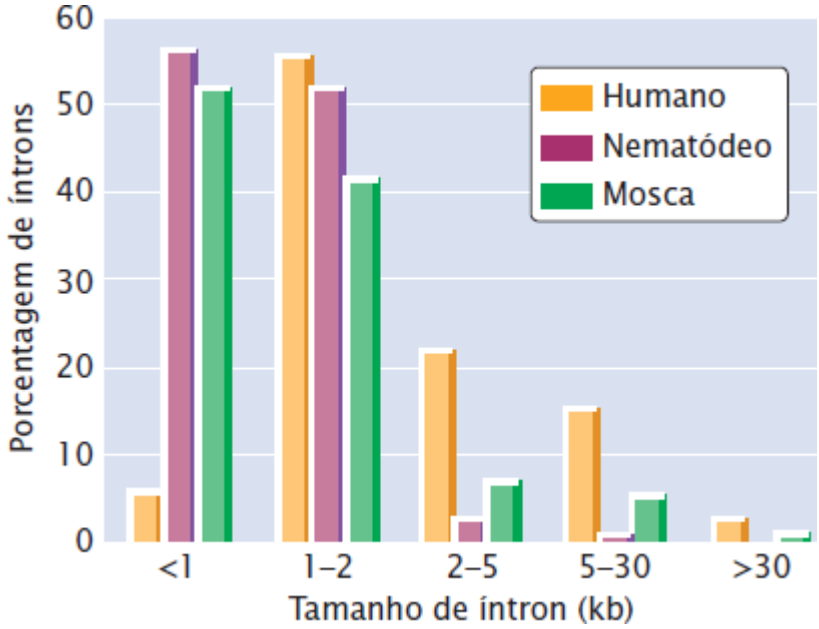


Figura 20.16 Os íntrons dos genes nos humanos são em geral mais longos que os íntrons dos genes nos nematódeos e moscas.

20.4 A proteômica analisa o conjunto complexo de proteínas encontrado em uma célula

Os dados da sequência do DNA oferecem muitas informações sobre a biologia de um organismo, mas elas não são toda a história. Muitos genes codificam proteínas e as proteínas executam a grande maioria das reações bioquímicas que modelam o fenótipo de um organismo. Embora as proteínas sejam codificadas pelo DNA, muitas proteínas sofrem modificações após a tradução e, nos organismos eucariotos mais complexos, existem muito mais proteínas do que genes. Assim, nos últimos anos, os biólogos moleculares voltaram sua atenção para a análise de conteúdo de proteína das células. O objetivo principal é determinar o proteoma, o conjunto completo de proteínas encontrado em uma célula. O estudo do proteoma é chamado **proteômica**.

Existem planos para identificar e caracterizar todas as proteínas no corpo humano, um esforço chamado de **Projeto Proteoma Humano**. O projeto catalogaria quais proteínas são encontradas em quais tipos de célula, onde cada proteína está localizada no interior da célula e quais outras proteínas interagem com estas. Muitos pesquisadores acreditam que esta informação será muito benéfica para identificar os alvos de fármacos, compreender a base biológica da doença e compreender a base molecular de muitos processos biológicos.

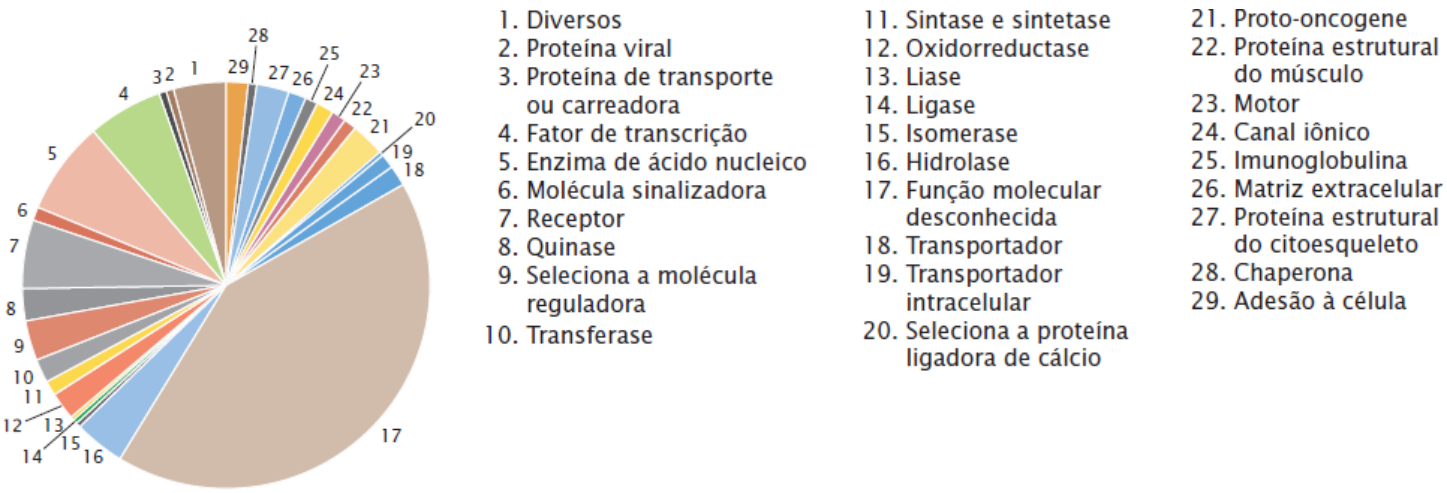


Figura 20.17 As funções para muitos genes humanos ainda precisam ser determinadas. A proporção do círculo ocupado por cada cor representa a proporção de genes que afetam várias funções conhecidas e desconhecidas.

Determinação das proteínas celulares

O procedimento básico para caracterizar o proteoma é primeiro separar as proteínas encontradas em uma célula e então identificar e quantificar as proteínas individuais. Um método para separar proteínas é a **eletroforese bidimensional em gel de poliacrilamida** (2D-PAGE), no qual as proteínas são separadas em uma dimensão pela carga elétrica, separadas em uma segunda dimensão pela massa e então coradas (**Figura 20.18 A**). Este procedimento separa as diferentes proteínas em manchas, com o tamanho de cada mancha proporcional à quantidade de proteína presente. Um típico gel 2D-PAGE pode ter muitas centenas a milhares de manchas (**Figura 20.18 B**).

Como o 2D-PAGE não detecta algumas proteínas em pequena quantidade e é difícil automatizar, os pesquisadores estão usando a cromatografia líquida para separar as proteínas. Na cromatografia líquida, uma mistura de moléculas é dissolvida em um líquido e passada por uma coluna preenchida com partículas sólidas. As diferentes afinidades para as fases líquida e sólida fazem com que alguns componentes da mistura viajem pela coluna mais lentamente do que outros, resultando na separação dos componentes da mistura.

O método tradicional para identificar uma proteína é remover seus aminoácidos um por vez e determinar a identidade de cada aminoácido removido. Esse método é lento demais e trabalhoso para analisar os milhares de proteínas existentes em uma célula. Atualmente os pesquisadores usam a **espectrometria de massa**, que é um método para determinar com precisão a massa molecular de uma molécula. Na espectrometria de massa, uma molécula é ionizada e é determinada sua taxa de migração em um campo elétrico. Como pequenas moléculas migram mais rapidamente do que as maiores, a taxa de migração consegue determinar com exatidão a massa da molécula.

Para analisar as proteínas com espectrometria de massa, primeiro uma proteína é digerida com a enzima tripsina, que rompe a proteína em fragmentos menores de peptídio, cada um com vários aminoácidos (**Figura 20.19 A**). A espectrometria de massa é usada para separar os peptídios com base em sua razão massa/carga elétrica (**Figura 20.19 B**). Esta separação produz um perfil de picos, com cada pico correspondendo a razão massa/carga elétrica de um peptídio (**Figura 20.19 C**). Um programa de computador procura em um banco de dados de proteínas para encontrar uma combinação entre o perfil gerado e o perfil esperado de uma proteína conhecida (**Figura 20.19 D**), permitindo que a proteína na amostra seja identificada. Usando bioinformática, o computador cria “produtos de digestão virtuais” e prevê os perfis de todas as proteínas encontradas em um genoma, dando as sequências de DNA dos genes codificadores de proteínas.

Métodos de espectrometria de massa também podem ser usados para medir a quantidade de cada proteína identificada. Com os avanços recentes, os pesquisadores agora fazem proteômica pela estratégia *shotgun*, que elimina a maioria dos estágios iniciais de separação de proteínas. Neste procedimento, uma mistura complexa de proteínas (como as existentes em uma amostra de tecido) é digerida e analisada com espectrometria de massa. A seguir, o programa de computador separa as proteínas existentes na amostra original a partir dos perfis de proteínas.

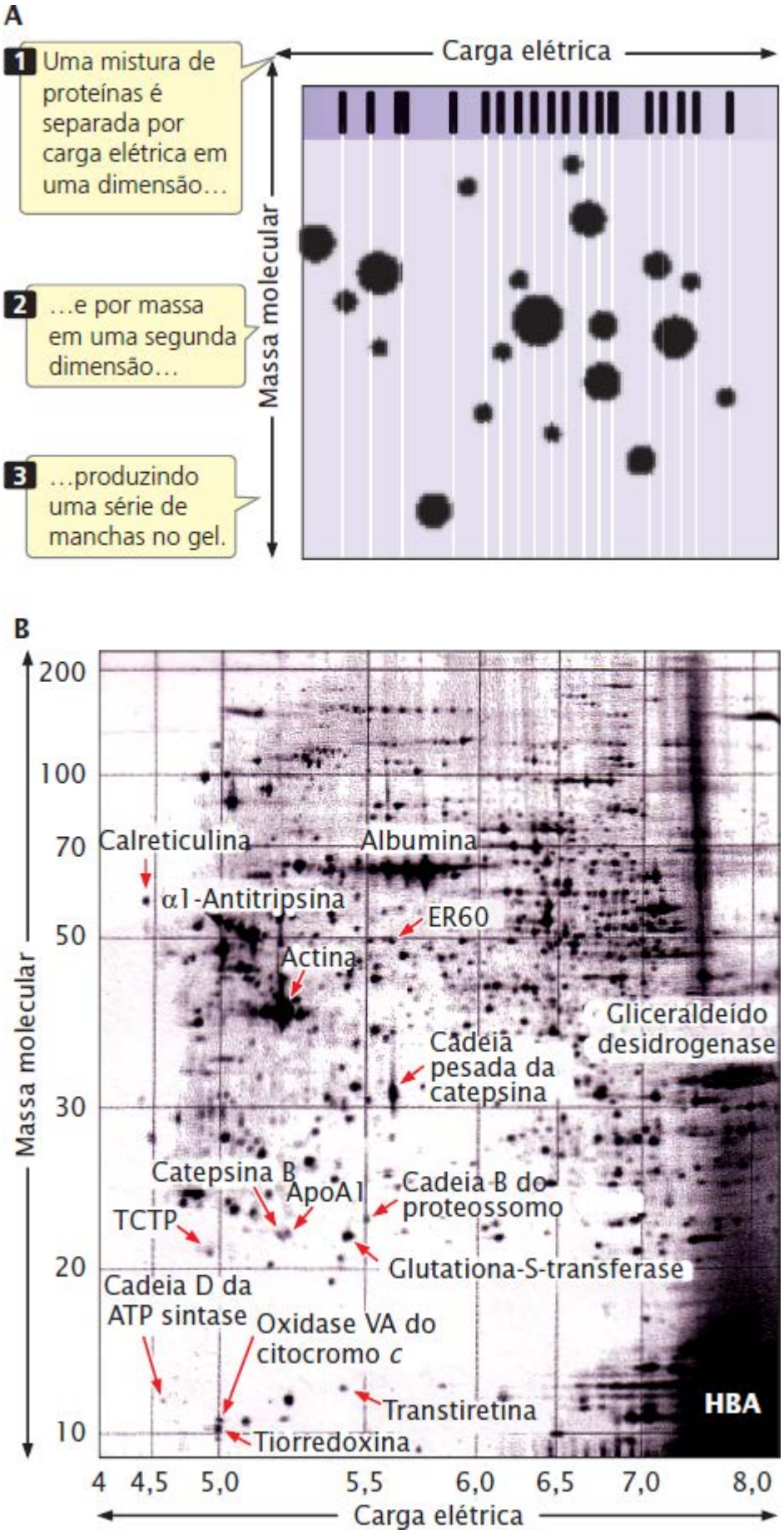
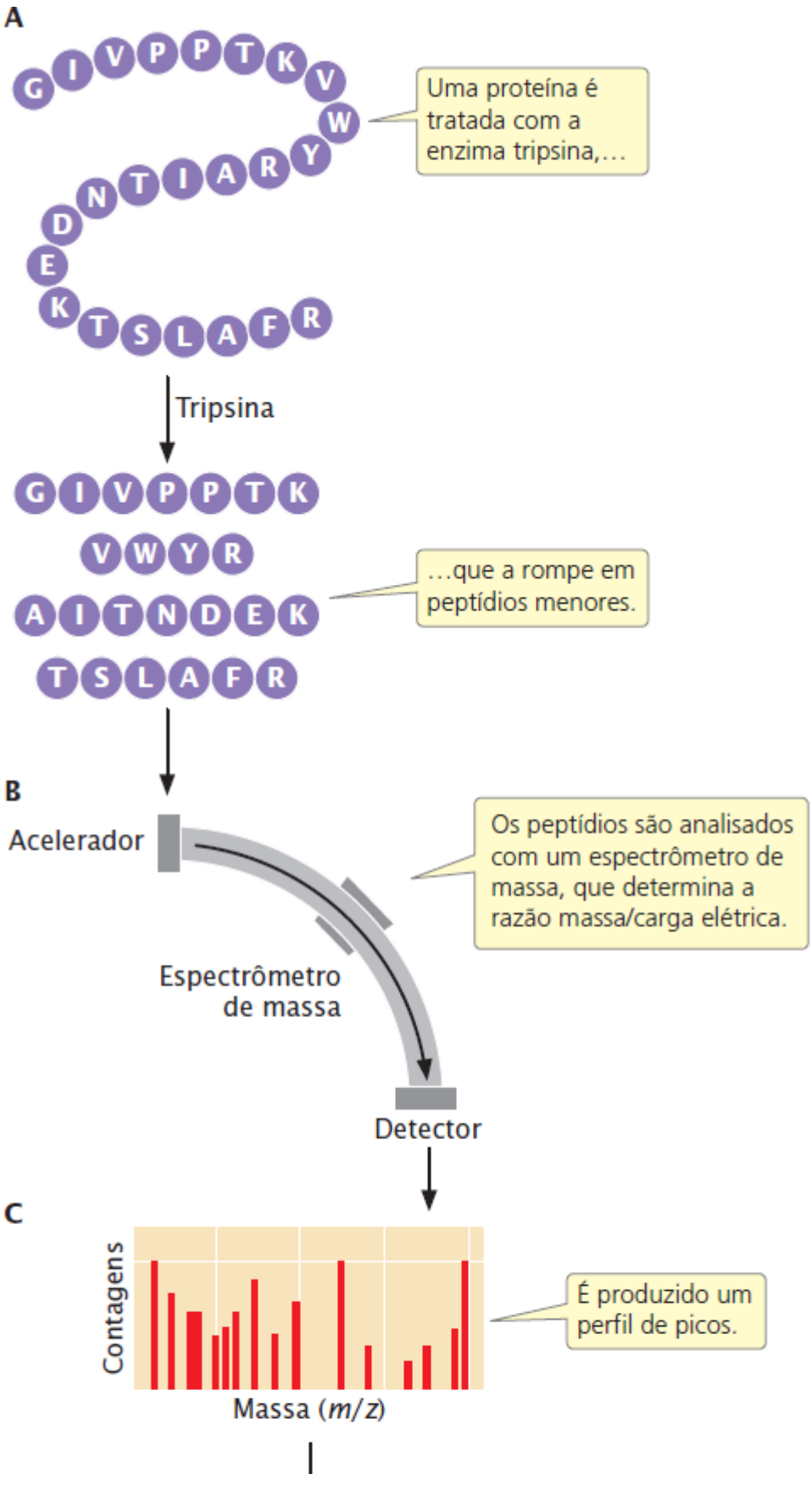


Figura 20.18 A eletroforese bidimensional em gel de acrilamida (2D-PAGE) pode ser usada para separar as proteínas celulares. (De G. Gibson and S. Muse. 2004. A Primer of Genome Science, 2e. Sinauer Associates, Inc. p. 274, Fig. 5.4.)

Mary Lipton e colaboradores usaram esta abordagem para estudar o proteoma da *Deinococcus radiodurans*, uma bactéria excepcional capaz de resistir a altas doses de radiação ionizante que são letais para todos os outros organismos. O genoma da *D. radiodurans* também já foi sequenciado. Lipton e colaboradores extraíram proteínas das bactérias, as digeriram com tripsina,

separaram os fragmentos com cromatografia líquida e determinaram as proteínas a partir dos fragmentos de peptídios com a espectrometria de massa. Eles identificaram 1.910 proteínas, que representam mais de 60% das proteínas previstas com base na sequência genômica.

Decifrar o proteoma de até mesmo uma única célula é um desafio. Cada célula tem uma sequência completa de genes, mas diferentes células expressam proteínas muito diferentes. Cada gene produz várias proteínas diferentes por meio de processamento alternativo (ver *Capítulo 14*) e processamento pós-tradução das proteínas (ver *Capítulo 15*). Uma típica célula humana contém até 100.000 diferentes proteínas que variam muito em abundância e nenhuma técnica como a PCR pode ser usada para amplificar as proteínas.



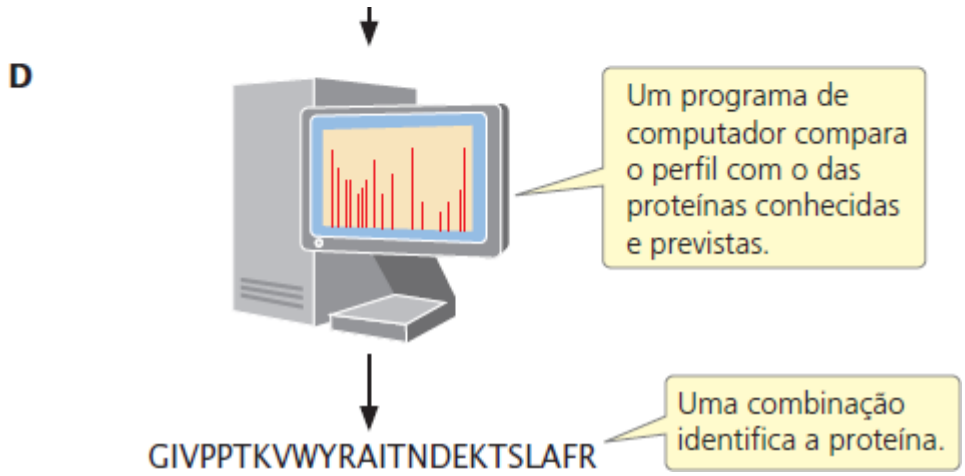


Figura 20.19 **A espectrometria de massa é usada para identificar as proteínas.**

Captura por afinidade

A proteômica visa não apenas identificar todas as proteínas em uma célula, mas também compreender como estas proteínas interagem e como sua expressão varia com o tempo. Os pesquisadores elaboraram várias técnicas para identificar as proteínas que interagem no interior da célula. Na **captura por afinidade**, um anticorpo (ver Figura 22.15) para uma proteína específica é usado para capturar uma proteína de uma mistura complexa de proteínas. A proteína capturada "levará" consigo qualquer proteína com a qual a proteína capturada interaja fisicamente. A mistura de proteínas capturadas pode, então, ser analisada por espectrometria de massa para identificar as proteínas. Várias modificações da técnica de captura por afinidade e outras técnicas podem ser usadas para determinar o conjunto completo de interações de proteínas em uma célula, chamado de **interatoma**.

Microarranjo de proteínas

As interações proteína-proteína também podem ser analisadas com **microarranjos de proteínas** (**Figura 20.20**), que são semelhantes aos microarranjos usados para examinar a expressão gênica. Com esta técnica, várias proteínas diferentes são aplicadas a uma lâmina de vidro como uma série de pontos, com cada ponto contendo uma proteína diferente. Em uma aplicação, cada ponto é um anticorpo para uma diferente proteína, marcado com uma etiqueta que fluoresce quando ligada. Um extrato do tecido é aplicado ao microarranjo de proteínas. Um ponto de fluorescência aparece quando uma proteína no extrato se liga a um anticorpo, indicando a existência dessa proteína no tecido.

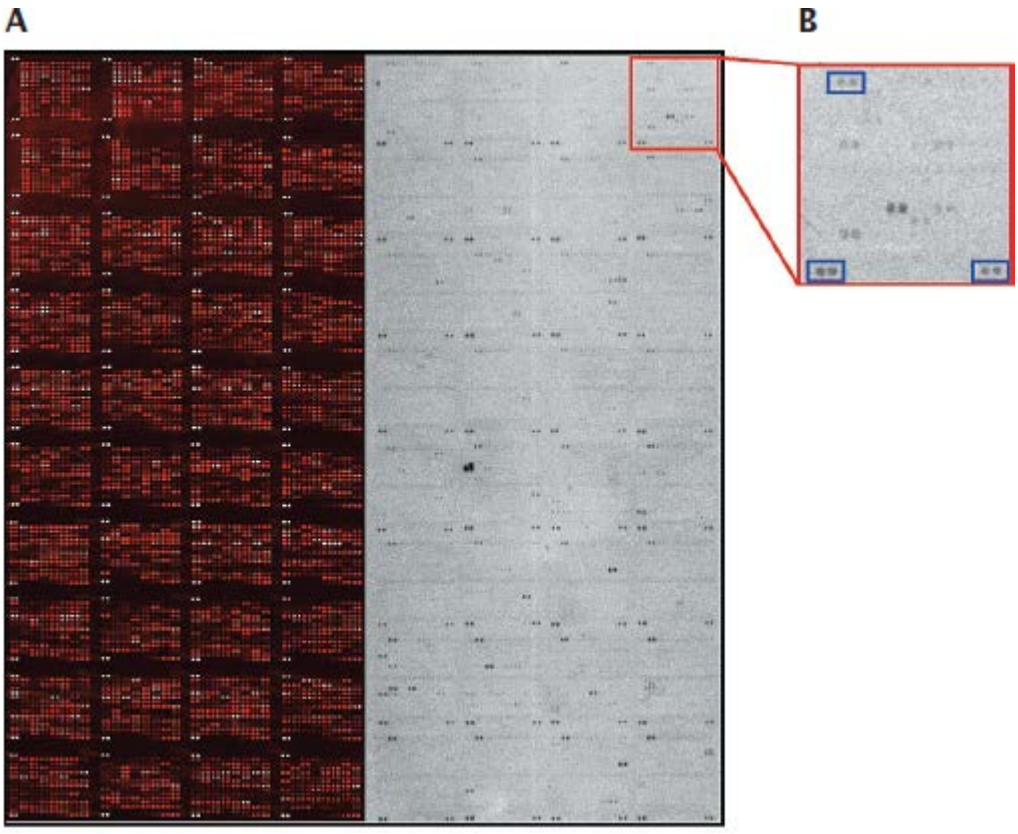


Figura 20.20 Os microarranjos de proteínas podem ser usados para examinar as interações das proteínas. A. Um microarranjo com 4.400 proteínas encontradas na levedura. **B.** O arranjo foi sondado com uma enzima que fosforila proteínas para determinar quais servem como substrato para essa proteína. Os pontos escuros representam proteínas que foram fosforiladas pela enzima. As proteínas que se fosforilam (autofosforilam) estão incluídas em cada bloco do microarranjo (apresentado nos boxes em azul) para servir como pontos de referência. (De D. Hall, J. Ptacek, and M. Snyder, 2006. Mechanisms of Ageing and Development 128 [2007] 161-167. © 2006, com permissão de Elsevier.)

Proteômica estrutural

A estrutura de alta resolução de uma proteína fornece muitas informações úteis. É uma fonte de esclarecimento sobre a função de uma proteína desconhecida, também pode sugerir a localização dos sítios ativos e fornecer informações sobre outras moléculas que interagem com a proteína. O conhecimento da estrutura de uma proteína sugere alvos para potenciais fármacos que possam interagir com a proteína. Como a estrutura fornece informações sobre a função, um dos objetivos da proteômica é determinar a estrutura de todas as proteínas encontradas em uma célula.

Dois procedimentos são usados atualmente para solucionar as estruturas de proteínas complexas: (1) cristalografia por raios X, na qual os cristais da proteína são bombardeados com raios X e os padrões de difração dos raios X são usados para determinar a estrutura (ver *Capítulo 10*) e (2) ressonância magnética (RM), que fornece informações sobre a posição de átomos específicos de uma molécula ao usar as propriedades magnéticas dos núcleos.

A cristalografia por raios X e a RM precisam da intervenção humana em muitos estágios e são muito lentas para determinar a estrutura de milhares de proteínas que podem existir dentro de uma célula. Como são necessárias estruturas de centenas de milhares de proteínas para os estudos da proteoma, os pesquisadores esperam ser capazes de prever a estrutura de uma proteína a partir de sua sequência de aminoácidos. Este método não é possível no momento, mas a esperança é que, se forem solucionadas estruturas de alta resolução em quantidade adequada, pode ser possível modelar a estrutura a partir apenas da sequência de aminoácidos. À medida que os cientistas trabalham em métodos automatizados que acelerem a determinação estrutural das proteínas, os profissionais da bioinformática estão desenvolvendo programas de computador melhores para prever a estrutura da proteína a partir de sua sequência.

Conceitos

O proteoma é o conjunto complexo de proteínas encontrado em uma célula. São usadas as técnicas de separação de proteínas e espectrometria de massa para identificar as proteínas existentes no interior de uma célula. As técnicas de captura por afinidade e microarranjos são usadas para determinar os conjuntos de proteínas que interagem entre si. A proteômica estrutural tenta determinar a estrutura de todas as proteínas.

Checagem dos conceitos 9

Por que conhecer a estrutura de uma proteína é importante?

Resumo dos conceitos

- A genômica é o campo da genética que tenta compreender o conteúdo, a organização e a função das informações genéticas contidas em genomas inteiros
- Os mapas genéticos posicionam os genes em relação a outros genes ao determinar as taxas de recombinação e são medidos em porcentagem de recombinação. Os mapas físicos baseiam-se nas distâncias físicas entre os genes e são medidos em pares de bases
- O Projeto Genoma Humano foi um esforço para determinar a sequência inteira do genoma humano. O projeto começou oficialmente em 1990; esboços preliminares da sequência do genoma humano foram completados em 2000. O resultado final do genoma humano foi completado em 2003
- O sequenciamento de um genoma inteiro exige a clivagem dele em pequenos fragmentos sobrepostos cujas sequências de DNA podem ser determinadas em reações de sequenciamento. As sequências individuais podem ser ordenadas em uma sequência de genoma inteiro usando uma abordagem baseada no mapa, na qual os fragmentos são montados em ordem ao usar mapas genéticos e físicos previamente criados ou com o uso de uma abordagem do genoma completo pela estratégia *shotgun*, no qual é usada a sobreposição entre os fragmentos para montá-los em uma sequência de genoma inteiro. Hoje em dia, praticamente todos os genomas são sequenciados usando o método *shotgun* do genoma completo
- Os polimorfismos de único nucleotídio são diferenças de base única no DNA entre organismos e são importantes marcadores nos estudos de ligação. Os organismos também podem apresentar diferenças no número de cópias de sequências de DNA, as chamadas variações no número de cópias

- Os sítios de sequências marcadas são sequências singulares de DNA cuja localização cromossômica foi determinada. As etiquetas de sequência expressa são marcadores associados a sequências de DNA expressas (transcritas) e podem ser usadas para encontrar os genes expressos em um genoma
- A bioinformática é uma síntese da biologia molecular e da ciência computacional para desenvolver ferramentas para armazenar, restaurar e analisar dados de sequências de DNA, cDNA e proteínas
- A metagenômica estuda os genomas de grupos inteiros de organismos. A biologia sintética é o desenvolvimento de técnica para criar genomas e organismos
- Os genes homólogos são correlatos em termos de evolução. Os ortólogos são as sequências homólogas encontradas em diferentes espécies, enquanto os parálogos são sequências homólogas na mesma espécie. A função gênica pode ser determinada pelo exame de sequências homólogas (tanto em ortólogos como em parálogos), cuja função já tenha sido determinada
- Um microarranjo é composto por fragmentos de DNA fixados em um padrão organizado em um suporte sólido, como filtro de náilon ou lâmina de vidro. Quando uma solução contendo uma mistura de DNA ou RNA é aplicada ao arranjo, qualquer ácido nucleico que é complementar à sonda usada se liga a esta. Os microarranjos podem ser usados para monitorar a expressão de milhares de genes simultaneamente
- Ao ligar uma sequência repórter com as sequências regulatórias de um gene, pode ser observado o padrão de expressão gênica ao procurar pelo produto da sequência repórter. Os genes que afetam uma função específica também podem ser identificados por meio de triagem de mutagênese de genoma inteiro
- A maioria das espécies procarióticas têm entre 1 e 3 milhões de pares de bases de DNA e de 1.000 a 2.000 genes. Em comparação com a dos genomas eucarióticos, a densidade gênica nos genomas procarióticos é relativamente uniforme, com cerca de um gene por 1.000 pb. Existe relativamente pouco DNA não codificador nos genes procarióticos. A transferência horizontal de genes (o movimento de genes entre diferentes espécies) foi um importante processo evolutivo nos procariotos
- Os genomas eucarióticos são maiores e seu tamanho varia mais do que os genomas procarióticos. Não existe relação clara entre a complexidade do organismo e a quantidade de DNA ou o número de genes nos organismos multicelulares. A maior parte dos genomas nos organismos eucarióticos é composta por DNA repetitivo. Os elementos de transposição são muito comuns na maioria dos genomas eucarióticos
- A proteômica determina o conteúdo de proteína de uma célula e as funções destas proteínas. As proteínas intracelulares podem ser separadas e identificadas com o uso de espectrometria de massa. A proteômica estrutural tenta determinar a forma tridimensional das proteínas.

Termos importantes

Bioinformática
Captura por afinidade
Contig
Desequilíbrio de ligação
Deserto gênico
Domínio da proteína
Duplicação segmentar
Eletroforese bidimensional em gel de poliacrilamida (2D-PAGE)
Espectrometria de massa
Estudo de associação genômica ampla
Etiqueta de sequência expressa (EST)
Família de multigenes
Genes homólogos
Genes ortólogos
Genes parálogos
Genômica
Genômica comparativa
Genômica estrutural
Genômica funcional
Haplótipo
Interatoma
Mapa físico
Mapa genético (ligação)

Metagenômica
Microarranjo
Microarranjo de proteínas
Microbioma
Notação
Polimorfismo de único nucleotídio (SNP)
Polimorfismo de único nucleotídio marcado (*tag*-SNP)
Projeto Proteoma Humano
Proteoma
Proteômica
Sequenciamento baseado no mapa
Sequenciamento *shotgun* do genoma completo
Sítio de sequências marcadas (STS)
Triagem de mutagênese
Transcriptoma
Variação no número de cópias (CNV)

Respostas da Checagem dos conceitos

1. Exatidão e resolução.
2. b.
3. Para catalogar e mapear SNPs e outras variantes genéticas humanas.
4. a.
5. d.
6. c.
7. As espécies com genomas maiores em geral têm mais genes que as espécies com genomas menores, então a densidade gênica é bem constante.
8. a.
9. A estrutura fornece, com frequência, informações importantes sobre como uma proteína funciona e os tipos de proteínas com as quais é provável que interaja.

Problema desenvolvido

Problema

Um segmento linear de DNA com 30 kb de comprimento é cortado primeiro com *Bam*HI, então com *Hpa*II e finalmente com *Bam*HI e *Hpa*II juntos. Foram obtidos os seguintes fragmentos a partir desta reação:

*Bam*HI: fragmentos de 20 kb, 6 kb e 4 kb
*Hpa*II: fragmentos de 21 kb e 9 kb
*Bam*HI e *Hpa*II: fragmentos de 20 kb, 5 kb, 4 kb e 1 kb

Desenhe um mapa de restrição do segmento de 30 kb de DNA, indicando as localizações dos sítios de restrição *Bam*HI e *Hpa*II.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Um mapa que inclua o número e as localizações relativas dos sítios de restrição para *Bam*HI e *Hpa*II e as distâncias em pb entre os sítios.

Quais informações são fornecidas para solucionar o problema?
- O segmento de DNA tem 30 kb de comprimento
- Os tamanhos dos fragmentos produzidos quando o DNA é cortado com *Bam*HI, com *Hpa*II e ambas enzimas juntas.

Para a solução deste problema, revise:

Mapas físicos na Seção 20.1 junto com o Problema desenvolvido na Seção 20.1.

Etapas para a solução

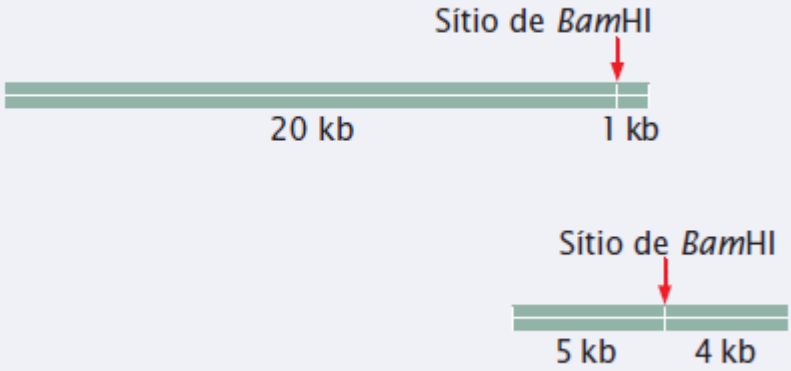
Quando cortado pela *Bam*HI sozinha, o segmento linear de DNA é rompido em três fragmentos, então existem obrigatoriamente dois sítios de restrição para *Bam*HI. Quando cortado com *Hpa*II sozinha, o clone do mesmo segmento de DNA é rompido em apenas dois fragmentos, então existe um único sítio para *Hpa*II.

Nota: Este problema pode ser solucionado corretamente por várias abordagens, esta solução aplica uma possível abordagem.

Dica: Para o DNA linear, o número de sítios de restrição é menos um que o número de fragmentos produzidos.

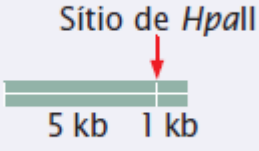
Vamos começar a determinar a localização destes sítios ao examinar os fragmentos de *Hpa*II. Observe que o fragmento de 21 kb produzido quando o DNA é cortado pela *Hpa*II não é encontrado nos fragmentos produzidos quando o DNA é cortado por *Bam*HI e *Hap*II juntas (a dupla digestão); este resultado indica que o fragmento de 21 kb de *Hap*II está no sítio de *Bam*HI. Se examinarmos os fragmentos produzidos pela dupla digestão, veremos que os fragmentos de 20 kb e 1 kb somam 21 kb; então um sítio de *Bam*HI tem de estar 20 kb de uma extremidade e 1 kb da outra extremidade.

Dica: Procure por fragmentos na dupla digestão que somem o comprimento de um fragmento presente na digestão simples.



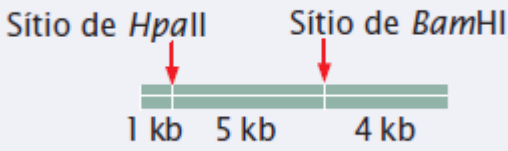
Da mesma maneira, observamos que o fragmento de 9 kb da *Hpa*II não aparece na dupla digestão e que os fragmentos de 5 kb e 4 kb desta digestão somam 9 kb; então outro sítio de *Bam*HI tem de estar a 5 kb de uma extremidade deste fragmento e 4 kb da outra extremidade.

Agora, vamos examinar os fragmentos produzidos quando o DNA é cortado pela *Bam*HI sozinha. Os fragmentos de 20 kb e 4 kb também são encontrados na dupla digestão, então nenhum destes fragmentos tem o sítio para *Hap*II. O fragmento de 6 kb, entretanto, não é encontrado na dupla digestão, e os fragmentos de 5 kb e 1 kb na dupla digestão somam 6 kb, então este fragmento tem um sítio para *Hap*II que está a 5 kb de uma extremidade e 1 kb da outra.



Solucionamos todos os sítios de restrição, mas ainda precisamos determinar a ordem dos sítios no

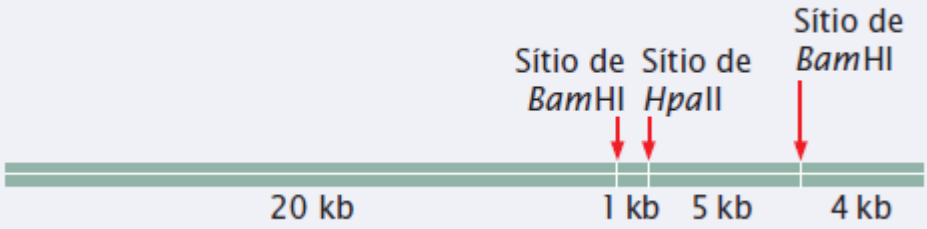
fragmento original de 30 kb.



Observe que o fragmento de 5 kb do sítio para *Hap*II precisa estar adjacente aos fragmentos de 1 kb e 4 kb; então ele obrigatoriamente está entre esses dois fragmentos.

Dica: Dois fragmentos na dupla digestão que foram produzidos ao cortar um fragmento na digestão simples estão necessariamente adjacentes um do outro.

Também estabelecemos que os fragmentos de 1 kb e 20 kb estão adjacentes, porque o fragmento de 5 kb está em um lado, o fragmento de 20 kb tem de estar no outro, completando o mapa de restrição.



Questões de compreensão

Seção 20.1

1. Qual é a diferença entre um mapa genético e um mapa físico? Qual tem melhor resolução e acurácia e por quê?
2. Qual é a diferença entre uma abordagem baseada no mapa para sequenciar um genoma inteiro e uma abordagem para genoma completo *shotgun*?
3. Quais são as questões éticas que surgem a partir das informações produzidas pelo Projeto Genoma Humano?
4. O que é polimorfismo de único nucleotídio (SNP)? Como os SNPs são usados nos estudos genômicos?
5. O que é um haplótipo? Como surgem diferentes haplótipos?
6. O que é desequilíbrio de ligação? Como ele gera os haplótipos?
7. Como é feito um estudo de associação genômica global?
8. O que é a variação no número de cópias? Como ela surge?
9. (a) O que é uma etiqueta de sequência expressa (EST)? (b) Como as ESTs são criadas? (c) Como as ESTs são usadas nos estudos genômicos?
10. Como os genes são reconhecidos nas sequências genômicas?

Seção 20.2

11. O que são sequências homólogas? Qual é a diferença entre ortólogos e parálogos?
12. Descreva vários métodos diferentes para deduzir a função de um gene ao examinar sua sequência de DNA.
13. O que é um microarranjo? Como ele pode ser usado para obter informações sobre a função do gene?
14. Explique como uma sequência repórter pode ser usada para fornecer informações sobre o padrão de expressão de um gene.
15. Explique resumidamente como é feita uma triagem de mutagênese.

Seção 20.3

16. Qual é a relação entre o tamanho do genoma e o número de genes nos procariotos?
17. O que é transferência horizontal de genes? Como ela é feita entre diferentes espécies de bactérias?

18. O conteúdo de DNA varia muito entre os diferentes organismos multicelulares. Esta variação está muito relacionada com o número de genes e a complexidade do organismo? Se não, o que explica algumas destas diferenças?
19. Mais da metade do genoma da *Arabidopsis thaliana* é composta por sequências duplicadas. Quais mecanismos podem ser responsáveis por estas substanciais duplicações?
20. O que é uma duplicação segmentar?
21. O genoma humano não codifica substancialmente mais domínios de proteínas do que os genomas invertebrados, mesmo assim ele codifica muito mais proteínas. Como são codificadas mais proteínas quando o número de domínios não difere muito?
22. (a) O que é genômica e qual é a diferença entre genômica estrutural e genômica funcional? (b) O que é genômica comparativa?

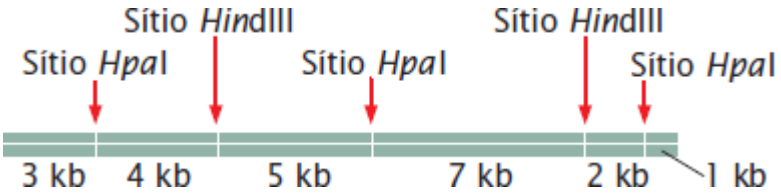
Seção 20.4

23. Qual a diferença entre proteômica e genômica?
24. Como a espectrometria de massa é usada para identificar as proteínas em uma célula?

Questões e problemas aplicados

Seção 20.1

*25. Um segmento de 22 kb de DNA tem os seguintes sítios de restrição:



Um lote deste DNA é primeiro totalmente digerido com *Hpa*I sozinha, então outro lote é totalmente digerido pela *Hin*dIII sozinha e finalmente, um terceiro lote é totalmente digerido pelas enzimas *Hpa*I e *Hin*dIII juntas. Os fragmentos resultantes de cada uma das três digestões são colocados em poços separados em um gel de agarose, separados por eletroforese em gel, e corado com brometo de etídio. Desenhe as bandas à medida que elas apareceriam no gel.

*26. Um segmento linear de DNA com 14 kb de comprimento é cortado primeiro pela *Eco*RI sozinha, depois por *Sma*I isoladamente e, por fim, por *Eco*RI e *Sma*I juntas. São obtidos os seguintes resultados:

Digestão pela *Eco*RI sozinha	**Digestão pela *Sma*I sozinha**	**Digestão por ambas *Eco*RI e *Sma*I**
Fragmento de 3 kb	Fragmento de 7 kb	Fragmento de 2 kb
Fragmento de 5 kb	Fragmento de 7 kb	Fragmento de 3 kb
Fragmento de 6 kb	-	Fragmento de 4 kb
-	-	Fragmento de 5 kb

Desenhe um mapa dos sítios de restrição de *Eco*RI e *Sma*I neste segmento de 14 kb de DNA, indicando as posições relativas dos sítios de restrição e as distâncias entre eles.

*27. A existência (+) ou não (–) de seis sítios de sequências marcadas (STSs) em cada um dos cinco clones (A-E) do cromossomo bacteriano artificial (BAC) é indicada no quadro a seguir. Usando estes marcadores, coloque os clones BAC na ordem correta e indique as localizações dos sítios STS neles.

	STSs					
Clone BAC	**1**	**2**	**3**	**4**	**5**	**6**
A	+	–	–	–	+	–
B	–	–	–	+	–	+
C	–	+	+	–	–	–

D	–	–	+	–	+	–
E	+	–	–	+	–	–

28. Um segmento linear de DNA foi dividido em fragmentos aleatórios sobrepostos e cada fragmento foi sequenciado. A sequência de cada fragmento é apresenta a seguir.

Fragmento 1:	5′-TAGTTAAAAC-3′
Fragmento 2:	5′-ACCGCAATACCCTAGTTAAA-3′
Fragmento 3:	5′-CCCTAGTTAAAAC-3′
Fragmento 4:	5′-ACCGCAATACCCTAGTT-3′
Fragmento 5:	5′-ACCGCAATACCCTAGTTAAA-3′
Fragmento 6:	5′-ATTTACCGCAAT-3′

Com base na sobreposição na sequência, crie uma sequência *contig* do fragmento original.

*29. Como a densidade dos genes encontrados no cromossomo 22 é comparada com a densidade dos genes encontrados no cromossomo 21, dois cromossomos de tamanhos semelhantes? Como o número de genes no cromossomo 22 é comparado com o número encontrado no cromossomo Y?

Para responder a estas perguntas, consulte o sítio eletrônico www.ensembl.org. No título *Species*, selecione *Human*. No lado esquerdo da próxima página clique em *Karyotype*. Aparecerão imagens dos cromossomos humanos. Clique no cromossomo 22 e selecione *Chromosome Summary*. Aparecerá uma imagem deste cromossomo e um histograma ilustrando as densidades dos genes totais (barras incolores) e dos genes conhecidos (barras coloridas). A parte inferior do diagrama informa o número total de genes, junto com o comprimento do cromossomo em pares de base. Escreva o comprimento total do cromossomo e o número de genes codificadores de proteínas.

Agora vá para o cromossomo 21 ao selecioná-lo a partir da lista de opções em Change Chromosome. Examine o comprimento total e o número de genes codificadores de proteínas para o cromossomo 21. Agora faça o mesmo para o cromossomo Y. Calcule a densidade gênica (número de genes/comprimento) para os cromossomos 22, 21 e Y.

a. Qual cromossomo tem a maior densidade e o maior número de genes? Qual tem a menor?
b. Examine com mais detalhes os genes na ponta do braço curto do cromossomo Y ao clicar na barra superior no histograma de genes. Pule para visualizar a localização. Será apresentada uma imagem mais detalhada. Quantos genes conhecidos são encontrados nesta região? Quantos genes codificadores de proteínas existem nesta região?

Cromossomos humanos 21 e 22. (Leonard Lessin/Science's Source.)

ANÁLISE DE DADOS

30. Nos últimos anos, as colônias de abelhas da América do Norte foram dizimadas pela rápida morte das abelhas operárias, uma condição chamada de distúrbio de colapso de colônia (CCD). Observada pela primeira vez pelos apicultores em 2004, esse distúrbio foi responsável pela perda de 50 a 90% das operações de apicultura nos EUA. Existem indícios que o CCD foi causado por um patógeno. Diana Cox-Foster e colaboradores (2007. *Science* 318:283-287) usaram uma abordagem metagenômica para tentar identificar o agente causador do CCD ao isolar o DNA de colmeias normais e das colmeias que tiveram CCD. Foram identificadas várias bactérias, fungos e vírus diferentes na análise metagenômica. O quadro a seguir fornece a porcentagem de colmeias com CCD e sem CCD que apresentaram teste positivo para quatro potenciais patógenos identificados na análise metagenômica. Com base nestes dados, qual potencial patógeno é o mais

provável de ser responsável pela CCD? Explique sua justificativa. Estes dados comprovam que este patógeno é a causa do CCD? Explique.

Vírus	Colmeias CCD infectadas (n = 30)	Colmeias sem CCD infectadas (n = 21)
Vírus israelense da paralisia aguda	83,3%	4,8%
Vírus da abelha Kashmir	100%	76,2%
Nosema apis	90%	47,6%
Nosema cernae	100%	80,8%

31. Noonan e colaboradores (2005. *Science* 309:597-599) começaram a estudar as sequências de genoma de uma espécie extinta de urso das cavernas. Eles extraíram o DNA de ossos de 40.000 anos de um urso das cavernas e usaram uma abordagem metagenômica para isolar, identificar e sequenciar o DNA deste urso. Por que eles usaram uma abordagem metagenômica quando seu objetivo era sequenciar o genoma de uma espécie (o urso das cavernas)?

Seção 20.2

32. Na Figura 20.8, explique porque os genes A2 e B2 são ortólogos e não parálogos.

*33. Examine a Figura 26.18. Os genes épsilon (e) e beta (b) no cromossomo 11 são ortólogos ou parálogos? Explique sua resposta.

*34. Os microarranjos podem ser usados para determinar os níveis de expressão gênica. Em um tipo de microarranjo, a hibridização dos cDNA vermelhos (câncer) e verdes (não câncer) é proporcional às quantidades relativas de mRNA nas amostras. O vermelho indica a expressão aumentada de um gene e o verde indica a expressão reduzida de um gene nas células experimentais em relação às células controle, o amarelo indica expressão igual nas células experimentais e controle e ausência de cor indica ausência de expressão em ambas as células.

Em um experimento, o mRNA de uma cepa de bactérias resistentes a antibiótico (células experimentais) é convertido em cDNA e marcado com nucleotídios vermelhos fluorescentes, o mRNA de uma cepa não resistente da mesma bactéria (células controle) é convertido em cDNA e marcado com nucleotídios verdes fluorescentes. Os cDNAs provenientes das células resistentes e não resistentes são misturados e hibridizados para um *chip* com pontos de DNA oriundo dos genes 1 ao 25. Os resultados obtidos são apresentados na ilustração desta questão. Quais conclusões você pode tirar sobre quais genes poderiam estar envolvidos na resistência aos antibióticos nestas bactérias? Como esta informação pode ser usada para projetar novos antibióticos que sejam menos vulneráveis à resistência?

35. Para os genes no microarranjo apresentado na parte inferior da Figura 20.9, a maioria destes genes é expressada de forma aumentada ou reduzida nos tumores de pacientes que ficaram sem câncer durante pelo menos 5 anos? Explique sua justificativa.

36. O que a fotografia na Figura 20.11 revela sobre a expressão da β-tubulina?

Seção 20.3

37. *Dictyostelium discoideum* é uma ameba social que vive no solo: boa parte do tempo, o organismo é composto por células

únicas, solitárias, mas, durante os períodos de fome extrema, as amebas se unem para formar agregados que têm muitas características dos organismos multicelulares. Os biólogos discutem se a *D. discoideum* é um organismo unicelular ou multicelular; seu genoma foi completamente sequenciado em 2005. O quadro a seguir lista algumas das características da *D. discoideum* e outros eucariotos (L. Eichinger *et al*. 2005. *Nature* 435:43-57).

a. Com base nos organismos listados no quadro além da *D. discoideum*, quais são as diferenças nas características do genoma entre os organismos unicelulares e multicelulares?
b. Com base nestes dados, você acha que o genoma da *D. discoideum* é mais semelhante aos dos eucariotos unicelulares ou dos eucariotos multicelulares? Explique sua resposta.

Quadro para o Problema 37

Característica	*D. discoideum*	*P. falciparium*	*S. cerevisiae*	*A. thaliana*	*D. melanogaster*	*C. elegans*	*H. sapiens*
Organismo	Ameba	Parasita da malária	Levedura	Planta	Mosca-da-fruta	Nematódeo	Humano
Celularidade	?	Uni	Uni	Multi	Multi	Multi	Multi
Tamanho do genoma (milhões de pares de bases)	34	23	13	125	180	103	2.851
Número de genes	12.500	5.268	5.538	25.498	13.676	19.893	22.287
Comprimento médio do gene (pb)	1.756	2.534	1.428	2.036	1.997	2.991	27.000
Genes com íntrons (%)	69	54	5	79	38	5	85
Número médio de íntrons	1,9	2,6	1,0	5,4	4,0	5,0	8,1
Tamanho médio do íntron (pb)	146	179	nd*	170	nd*	270	3.365
Média de G + C (éxons)	27%	24%	28%	28%	55%	42%	45%

*nd = não determinado

*38. Como comparar as características genômicas a seguir dos organismos procarióticos em relação aos eucarióticos? Como eles se comparam entre os eucariotos?

a. Tamanho do genoma.
b. Número de genes.
c. Densidade gênica (pb/gene).
d. Número de éxons.

ANÁLISE DE DADOS

39. Um grupo de cientistas sequenciou os genomas de 12 espécies de *Drosophila* (*Drosophila* 12 Genomes Consortium. 2007.

Nature 450:203-218). O quadro a seguir apresenta os dados sobre tamanho do genoma e número de genes que codificam proteínas a partir deste estudo. Faça uma planilha com o número de genes codificadores de proteínas como uma função do tamanho do genoma para as 12 espécies de *Drosophila*. Existe uma relação entre o tamanho do genoma e o número de genes nas moscas-da-fruta? Como ele se compara com relação ao tamanho do genoma e número de genes sobre todos os eucariotos?

Características dos genomas das 12 espécies de *Drosophila*.

Espécie	Tamanho do genoma (milhões de pares de bases)	Número de genes codificadores de proteínas
D. melanogaster	200	13.733
D. simulans	162	15.983
D. sechellia	171	16.884
D. yakuba	190	16.423
D. erecta	135	15.324
D. ananassae	217	15.276
D. pseudoobscura	193	16.363
D. persimilis	193	17.325
D. willistoni	222	15.816
D. virilis	364	14.680
D. mojavensis	130	14.849
D. grimshawi	231	15.270

Seção 20.4

40. Um cientista determina os genomas e proteomas completos de uma célula do fígado e do músculo da mesma pessoa. Você esperaria grandes diferenças nos genomas ou nos proteomas destes dois tipos de células? Explique sua resposta.

Questões desafiadoras

Seção 20.1

41. O genoma da *Drosophila melanogaster*, mosca-da-fruta, foi sequenciado em 2000. Entretanto, esta sequência completa não incluiu a maioria das regiões de heterocromatina. A heterocromatina foi sequenciada em 2007 (R. A. Hoskins *et al.* 2007. *Science* 316:1625-1628). A maioria das sequências de genoma não inclui a heterocromatina. Por que a heterocromatina não costuma ser sequenciada em projetos de genoma? (Dica: Consulte o *Capítulo 11* para uma discussão mais detalhada sobre heterocromatina).
42. Nos estudos metagenômicos, uma comparação das sequências de RNA ribossômico é usada para determinar o número de espécies diferentes existentes. Quais são algumas das características das sequências ribossômicas que as tornam úteis para determinar as espécies existentes?

*43. Alguns biólogos sintéticos propuseram criar um organismo totalmente novo, de vida livre com um genoma mínimo, o menor conjunto de genes que permita a replicação do organismo em um ambiente específico. Este genoma poderia ser usado para projetar e criar, do “nada” organismos novos que poderiam desempenhar tarefas específicas como degradar materiais tóxicos no ambiente.
    a. Como poderia ser determinado o genoma mínimo necessário para vida?
    b. Quais poderiam ser, se houvesse alguma, as questões éticas e sociais associadas à criação de um organismo totalmente

novo com um genoma mínimo?

21 Epigenética

Como a dieta do seu avô pode afetar a sua saúde

A comunidade Överkalix, no norte da Suécia, localiza-se acima do Círculo Ártico, com um clima inóspito caracterizado por longos invernos frios e períodos curtos de verão. A comunidade sempre foi pequena; hoje em dia existem menos de 4.000 residentes. No século 19 e no início século 20, existiam poucas estradas na região, e o transporte durante o inverno era limitado pelo gelo e pela neve. Por causa da localização no extremo norte, era um desafio cultivar plantações. Quando as plantações morriam, a distância da região limitava a importação de alimento e as pessoas morriam de fome. As plantações fracassavam com frequência durante todo o século 19 (p. ex., em 1800, 1812, 1821 e 1829) e o período de 1831 a 1836 foi marcado por um fracasso total das plantações e extrema provação. Entretanto, em razão da imprevisibilidade do clima, com frequência anos de colheita bemsucedida e fartura de alimentos ocorreram após esses anos difíceis.

Nos anos de 1980, os pesquisadores se interessaram pelos efeitos da fartura e da escassez sobre a saúde duradoura das pessoas no norte da Suécia. Eles queriam determinar se os períodos de escassez de alimento que as pessoas passavam na infância afetavam a saúde dos seus descendentes. Após examinar cuidadosamente a estatística de colheitas, preços de grãos e outros fatos históricos, os pesquisadores conseguiram determinar a disponibilidade de alimentos na durante o século 19 e início do século 20. Eles também examinaram os prontuários dos habitantes, graças à disponibilidade de registros médicos centralizados que a Suécia mantém sobre todos seus cidadãos.

Os pesquisadores se concentraram na saúde de três grupos de pessoas, nascidas em 1890, 1905 e 1920. Examinaram o tempo de vida desses indivíduos e examinaram seu risco de morrer de doença cardiovascular e diabetes melito. Depois, rastrearam seus pais e avôs, determinaram a disponibilidade de alimento quando eram crianças e procuraram por correlações entre a dieta dos pais e avôs e a saúde de seus descendentes.

O que os pesquisadores encontraram foi impressionante. Os indivíduos cujos pais e avós haviam sido expostos a períodos de falta de alimentos quando crianças viveram mais do que os indivíduos cujos ancestrais haviam sido expostos a excesso de alimentos. Por outro lado, as pessoas cujos ancestrais cresceram durante os períodos de fartura de alimentos morreram mais cedo e eram mais predispostas a doenças cardiovasculares e diabetes melito. Por exemplo, se um avô paterno teve acesso a alimentos em excesso quando criança, a probabilidade foi quatro vezes maior de que os netos morressem de diabetes melito do que os netos de pessoas que não foram expostas a alimentos em excesso durante a infância.

Como a quantidade de alimento disponível para uma pessoa durante a infância pode afetar a saúde de seus filhos e netos que vivem 20 a 60 anos depois? Um dos princípios da genética moderna é que nossos genes são estáveis (exceto por raras mutações) e não são alterados pelo ambiente, então como a dieta pode influenciar os traços dos descendentes por duas gerações? O efeito da dieta do avô sobre as gerações seguintes foi especialmente incrível. As mães fornecem o citoplasma do óvulo a seus descendentes e um ambiente uterino, assim como os genes; mas os espermatozoides de seus pais contribuem apenas com um conjunto de genes paternos a seus descendentes.

Os pesquisadores propuseram que o efeito observado foi consequente à epigenética: mudanças na cromatina e no DNA que são hereditárias, mas não envolvem alteração na sequência de bases do DNA. A herança epigenética não foi prevista por Mendel, nem até recentemente, pela maioria dos geneticistas modernos, mas parece que os processos epigenéticos são importantes na herança de muitos genótipos. Atualmente, o estudo da epigenética é o foco de pesquisa intensiva.

Outra pergunta que esse estudo levanta é por que as condições adversas de fome durante a infância *reduzem* o risco de morrer de doença cardiovascular e diabetes melito nas gerações futuras, enquanto as condições de fartura de alimento *aumentam* o risco? Pode-se esperar justamente o oposto, que o estresse nutricional durante a infância aumente o risco de morrer, enquanto o excesso de alimento reduz. Os biólogos evolucionistas propuseram uma explicação para essa relação, que também foi observada em outros estudos. Essa explicação, chamada de *hipótese do fenótipo poupador*, baseia-se na suposição de que as informações sobre o ambiente onde os pais viveram pode ser útil para os descendentes, permitindo que eles respondam de formas que

aumentem sua própria sobrevida e reprodução. Essa hipótese propõe que quando as condições ambientais são ruins para os pais, é provável que elas persistam e também sejam ruins para os descendentes. Portanto, quando o genitor passa por períodos difíceis, a seleção natural favorece os genitores que produzem descendentes metabolicamente poupadores, descendentes que comam o máximo possível quando o alimento estiver disponível, minimizem o gasto de energia e acumulem calorias, porque o ambiente dos genitores prevê que os descendentes terão pouco alimento disponível. Essa estratégia foi provavelmente vantajosa no passado distante, antes da agricultura, mas frequentemente tem efeitos deletérios na sociedade moderna. A ingestão de muita comida, a minimização do gasto energético e o acúmulo de calorias quando há fartura de alimentos resultam, com frequência, em obesidade, cardiopatia e diabetes melito, como foi observado nas crianças e netos das pessoas de Överkalix.

Este capítulo estuda a epigenética, a explicação proposta para o efeito da dieta sobre a saúde dos residentes de Överkalix. Começamos discutindo a origem do termo epigenética e o que esse termo abrange hoje em dia. Então vamos revisar os tipos de mudanças que ocorrem e os principais processos que alteram a estrutura da cromatina. Vamos examinar como as mudanças na estrutura da cromatina podem ser transmitidas para as células e as gerações futuras. Depois analisaremos vários efeitos epigenéticos, incluindo paramutação, efeitos comportamentais, efeitos de substâncias químicas, efeitos metabólicos, efeitos nos gêmeos monozigóticos, inativação do X, diferenciação celular e *imprinting* genômico. E terminamos o capítulo discutindo os esforços para mapear a localização de genoma amplo dos marcos epigenéticos, o epigenoma.

21.1 O que é epigenética?

O termo epigenética foi usado pela primeira vez por Conrad Waddington em 1942 para descrever como, por meio do processo de desenvolvimento, um genótipo produz um fenótipo. Waddington combinou as palavras “epigênese” que é como um embrião se desenvolve, com “genética”, o estudo dos genes e da hereditariedade. O objetivo de Waddington era incentivar a união da genética e do desenvolvimento. Entretanto, seu uso do termo antecedeu nossa moderna compreensão do DNA e da estrutura do cromossomo e atualmente a epigenética tem um significado mais restrito.

A raiz grega *epi* significa *sobre* ou *acima*; a epigenética veio para representar a herança de variação acima e além das diferenças na sequência de DNA. Atualmente, a epigenética se refere, de modo geral, aos fenótipos e processos que são transmitidos para outras células e às vezes, a futuras gerações, mas não são o resultado das diferenças na sequência de bases do DNA. Os efeitos epigenéticos são causados por mudanças na expressão gênica resultantes de alterações à estrutura de cromatina ou outros aspectos da estrutura do DNA, como a metilação do DNA. Uma definição de um traço epigenético é: um fenótipo herdado com estabilidade resultante de mudanças na cromatina sem alterações na sequência de DNA. Alguns pesquisadores ampliaram a definição de epigenética de modo a incluir qualquer alteração da estrutura da cromatina ou do DNA que afete a expressão gênica. Aqui, usaremos a epigenética para se referir às mudanças na expressão gênica e/ou em um fenótipo que são potencialmente herdados sem alteração da sequência de bases de DNA envolvida.

Muitas mudanças epigenéticas são estáveis, persistindo por divisões celulares ou até gerações. Entretanto, as alterações epigenéticas também são potencialmente influenciadas por fatores ambientais. Por exemplo, foi demonstrado que o estresse ambiental altera a metilação do gene *Bdnf* do rato, que codifica um fator de crescimento importante no desenvolvimento do cérebro. A metilação do DNA foi correlacionada com a expressão de genes e os fenótipos produzidos. Como veremos, a metilação do DNA alterada consegue ser replicada durante a divisão celular, resultando em descendentes com o novo fenótipo, embora não exista diferença correspondente na sequência de bases de DNA dos indivíduos que “herdam” o novo fenótipo. O fato de que os traços epigenéticos podem ser induzidos por fatores ambientais e transmitidos a futuras gerações foi interpretado por alguns para significar que os genes têm memória por meio da epigenética, que os fatores ambientais que atuam sobre os indivíduos podem ter efeitos que são transmitidos para futuras gerações, como foi observado com o efeito da dieta na expectativa de vida na introdução deste capítulo. A epigenética tem sido chamada de “herança, mas não como a conhecemos”.

A epigenética está fornecendo uma explicação para como alterações fora da sequência de DNA conseguem influenciar o fenótipo e como essas mudanças são herdadas. Também está fazendo contribuições importantes para o estudo do comportamento, ciência ambiental, câncer, neurobiologia e farmacologia. Resolva o Problema 2

Conceitos

Os efeitos epigenéticos são fenótipos e processos transmitidos para outras células e, às vezes, a futuras gerações, mas não são o resultado das diferenças na sequência de bases do DNA. O estudo da epigenética está fazendo importantes contribuições para muitas áreas da Biologia.

21.2 Vários processos moleculares levam a mudanças epigenéticas

A epigenética altera a expressão dos genes; essas alterações são estáveis o suficiente para serem transmitidas por mitose (e, às vezes, por meiose), mas também podem ser alteradas. A maioria das evidências sugere que os efeitos genéticos são produzidos por mudanças físicas na estrutura da cromatina. No Capítulo 11, analisamos várias mudanças químicas no DNA e nas proteínas histona que afetam a estrutura da cromatina, incluindo metilação do DNA, modificação das proteínas histona e reposicionamento de nucleossomos. No Capítulo 17, discutimos o papel dessas alterações na expressão dos genes. Acredita-se que essas mudanças na estrutura da cromatina sejam importantes nos traços epigenéticos. O Capítulo 14 discutiu as pequenas moléculas de RNA, importantes para produzir os efeitos epigenéticos.

Vamos analisar três tipos de mecanismos moleculares que alteram a estrutura da cromatina e sustentam muitos fenótipos epigenéticos: (1) mudanças nos padrões de metilação do DNA; (2) modificações químicas das proteínas histona e (3) moléculas RNA que afetam a estrutura da cromatina e a expressão gênica.

Conceitos

Muitos fenótipos epigenéticos são o resultado de alterações na estrutura da cromatina, mediadas por três processos principais: metilação do DNA, modificação da histona e moléculas de RNA.

Metilação do DNA

O mecanismo mais bem compreendido da mudança epigenética é a metilação do DNA. A metilação do DNA refere-se ao acréscimo de grupos metila às bases de nucleotídios. Nos eucariotos, o tipo predominante de metilação de DNA é a metilação da citosina para produzir 5-metilcitosina (**Figura 21.1 A**). Como foi discutida no Capítulo 17, a metilação do DNA está associada à repressão da transcrição.

Ela ocorre, com frequência, nas bases citosina que estão imediatamente adjacentes aos nucleotídios guanina, chamados de dinucleotídios CpG (em que p representa o grupo fosfato que conecta os nucleotídios C e G). Nos dinucleotídios CpG, os nucleotídios citosina nas duas fitas de DNA estão em diagonal um em relação ao outro. Em geral, ambas as bases citosina estão metiladas, então os grupos metila surgem em ambas as fitas de DNA, como apresentado a seguir e na **Figura 21.1 B**.

$$5'-\overset{m}{C}G-3'$$
$$3'-G\underset{m}{C}-5'$$

Nas plantas, a metilação do DNA também ocorre em trinucleotídios CpNpG, em que N representa um nucleotídio com qualquer base.

Algumas regiões do DNA têm mais dinucleotídios CpG e são chamados de ilhas de CpG. Nas células de mamíferos, as ilhas de CpG estão localizadas nos promotores de genes ou próximas a estes. Essas ilhas de CpG não estão metiladas quando os genes são transcritos de forma ativa. Entretanto, a metilação das ilhas de CPG próximas de um gene leva a repressão da transcrição.

As células reprimem e ativam os genes ao metilar e desmetilar as bases citosina. Enzimas chamadas DNA metiltransferases metilam DNA ao adicionar grupos metila a bases de citosina para criar 5-metilcitosina. Outras enzimas chamadas de desmetilases removem os grupos metila, convertendo 5-metilcitosina de volta para citosina (ver *Capítulo 17*).

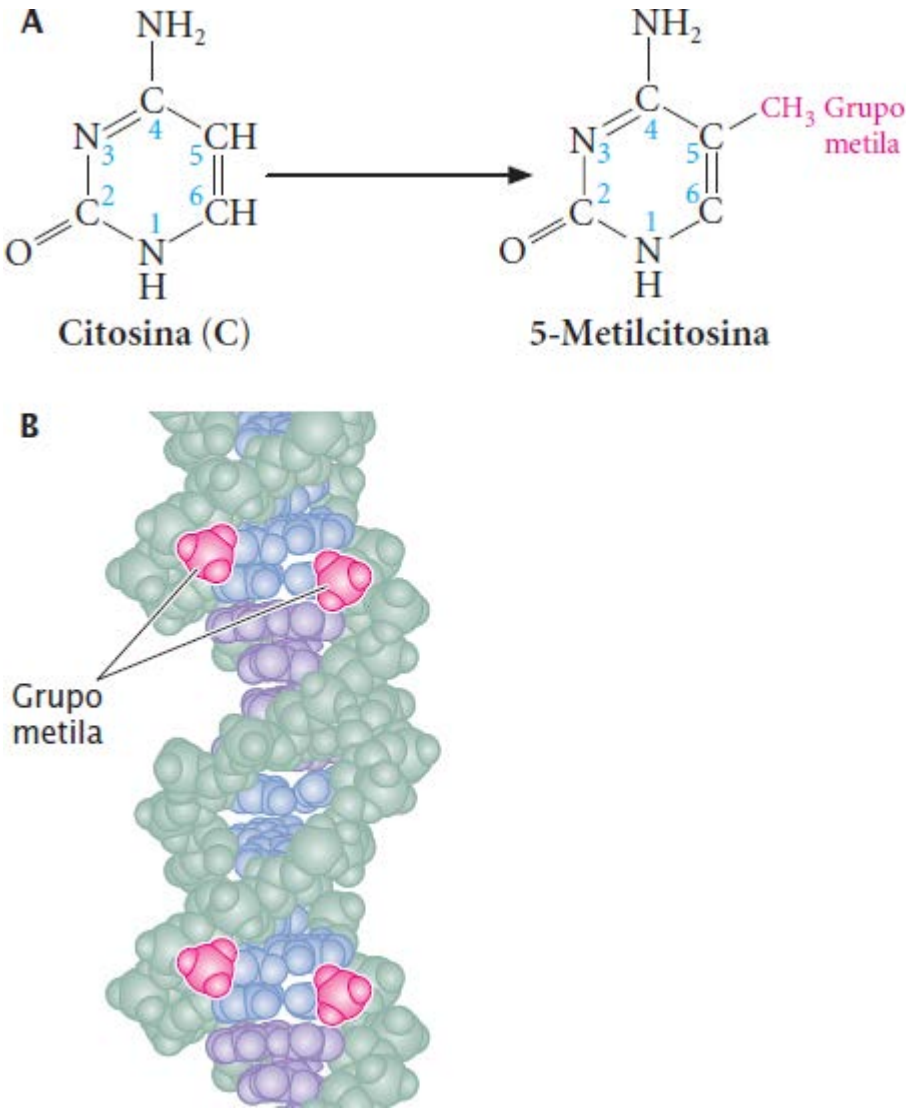


Figura 21.1 A metilação do DNA é uma modificação epigenética comum da cromatina. A. As bases citosina são modificadas para formar 5-metilcitosina. **B.** Estrutura tridimensional de DNA mostrando a metilação dos dinucleotídios CpG.

Manutenção da metilação. O fato de que as mudanças epigenéticas são transmitidas para outras células e (às vezes) para gerações futuras significa que as mudanças na estrutura da cromatina associadas aos fenótipos epigenéticos têm de ser fielmente mantidas quando os cromossomos se replicarem. Como as mudanças epigenéticas são mantidas e replicadas durante o processo de divisão celular?

A metilação de sequências CpG significa que duas bases citosina metiladas estão em diagonal uma da outra em fitas opostas. Antes da replicação, as bases citosina em ambas as fitas são metiladas (**Figura 21.2**). Imediatamente após a replicação semiconservativa, a base citosina na fita molde estará metilada, mas a base citosina na fita recém-replicada, não. Enzimas metiltransferases especiais reconhecem o estado hemimetilado dos dinucleotídios CpG e adicionam grupos metila a bases citosina não metiladas, criando duas novas moléculas de DNA totalmente metiladas. Dessa forma, o padrão de metilação do DNA é mantido durante a divisão celular. Resolva o Problema 26

Metilação de DNA nas abelhas. Um exemplo incrível de epigenética é observado nas abelhas. As abelhas-rainhas e as abelhas-operárias são fêmeas, mas a semelhança termina aqui. A rainha é maior e desenvolve ovários funcionais, enquanto as operárias são pequenas e estéreis (**Figura 21.3**). A rainha faz o voo do acasalamento e gasta sua via inteira se reproduzindo, enquanto as operárias passam a vida toda coletando néctar e pólen, supervisionando a rainha e criando seus descendentes. Apesar dessas significativas diferenças na anatomia, na fisiologia e no comportamento, as rainhas e operárias são geneticamente idênticas, evoluindo a partir de óvulos comuns. Sua diferença está na dieta: as abelhas-operárias produzem e alimentam algumas larvas fêmeas com uma substância especial chamada geleia real, que faz com que essas larvas se desenvolvam em rainhas. Outras larvas recebem nutrientes básicos e se tornam operárias. Essa simples diferença na dieta afeta muito a expressão gênica, fazendo com que diferentes genes sejam ativados nas abelhas e resultem em um conjunto muito diferente de traços fenotípicos.

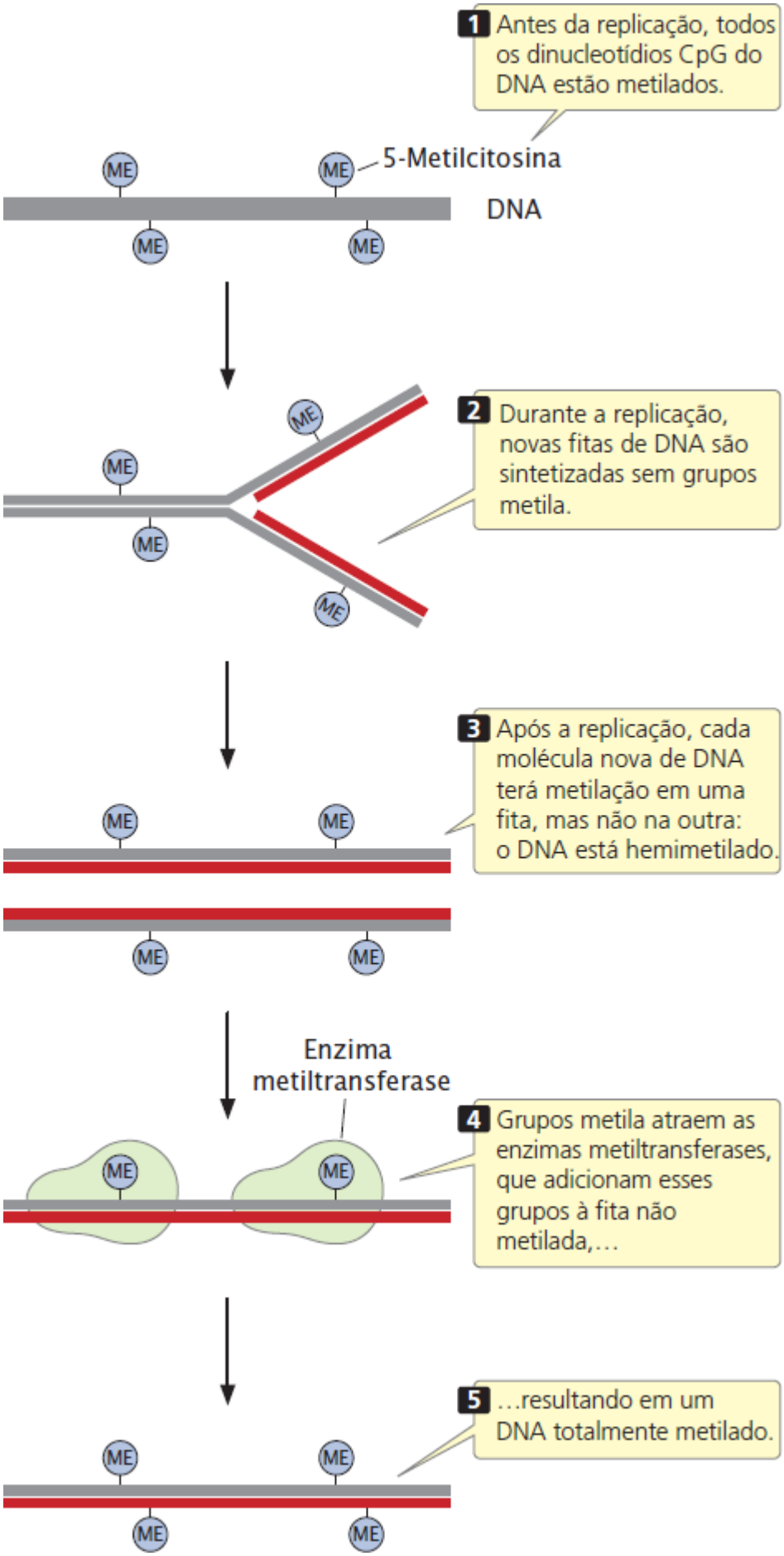


Figura 21.2 A metilação do DNA é mantida estável durante a replicação do DNA.

Figura 21.3 As mudanças epigenéticas são responsáveis por diferenças nos fenótipos das abelhas (*Apis mellifera*) rainha (*esquerda*) e operárias (*direita*). (WILDLIFE GmbH/Alamy.)

Como a geleia real afeta a expressão gênica é um mistério, mas a pesquisa atualmente sugere que ela muda um marcador epigenético. Em 2008, Ryszard Kucharski e seus colaboradores demonstraram que a geleia real silencia a expressão de um genechave chamado *Dnmt3*, cujo produto adiciona grupos metila ao DNA. Com o *Dnmt3* desligado, o DNA da abelha é menos metilado e muitos genes que normalmente estão silenciados nas operárias são expressos, levando ao desenvolvimento das características da rainha. Kucharski e seus colaboradores demonstraram a importância da metilação do DNA no desenvolvimento da rainha ao injetar pequenas moléculas de RNA nas larvas de abelhas (pequenos RNA de interferência, ou siRNAs, ver *Capítulo 14*) que inibem especificamente a expressão de *Dnmt3*. Essas larvas tinham níveis menores de metilação do DNA e muitas se desenvolveram como rainhas com ovários funcionais. Esse experimento demonstrou que a geleia real produz mudanças epigenéticas (menos a metilação do DNA), que são transmitidas pela divisão celular e modificam as vias do desenvolvimento, consequentemente produzindo uma abelharainha. Resolva o Problema 25

Repressão da transcrição pela metilação do DNA. Como a metilação do DNA suprime a expressão gênica? O grupo metila da 5-metilcitosina se localiza no sulco principal do DNA, que é reconhecido por muitas proteínas ligadoras de DNA. A localização do grupo metila no sulco principal inibe a ligação de fatores de transcrição e outras proteínas necessárias para que ocorra a transcrição. A 5-metilcitosina também atrai algumas proteínas que reprimem diretamente a transcrição. Além disso, a metilação do DNA atrai as enzimas histona desacetilase, que removem os grupos acetila das caudas das proteínas histona, alterando a estrutura da cromatina de forma que reprime a transcrição (ver *Capítulo 17*).

Conceitos

As bases citosina são metiladas para formar a 5-metilcitosina, que está associada à repressão da transcrição. A metilação do DNA é mantida estável durante a replicação pelas enzimas metiltransferases que reconhecem o estado hemimetilado dos nucleotídios CpG e adicionam grupos metila para as bases citosina não metiladas.

Checagem dos conceitos 1

Qual das seguintes afirmativas é verdadeira sobre as ilhas de CpG?

a. Elas estão metiladas próximas dos promotores de genes ativamente transcritos.
b. Elas não estão metiladas próximas dos promotores de genes ativamente transcritos.
c. A acetilação das ilhas de CpG leva à repressão da transcrição.
d. As ilhas de CpG codificam moléculas de RNA que ativam a transcrição.

Modificações da histona

As mudanças epigenéticas também podem ocorrer pela modificação das proteínas histona. Nas células eucarióticas, o DNA é complexado a proteínas histona na forma de nucleossomos, que são as unidades repetidas básicas da estrutura da cromatina (ver *Capítulo 11*). Já foram detectadas mais de 100 diferentes modificações pós-tradução das proteínas histona. Muitas dessas modificações ocorrem nas caudas com carga elétrica positiva das proteínas histona, que interagem com o DNA e afetam a

estrutura da cromatina. As modificações nas histonas incluem o acréscimo de fosfatos, grupos metila, grupos acetila e ubiquitina a suas caudas. Essas modificações alteram a estrutura da cromatina e afetam a transcrição dos genes (ver *Capítulo 17*). As modificações também servem como sítios de ligação para proteínas como fatores de transcrição necessários para esse processo.

A adição de grupos acetila a aminoácidos nas caudas da histona (acetilação da histona) em geral desestabiliza a estrutura da cromatina, fazendo com que ela assuma uma configuração mais aberta e esteja associada a mais transcrição (ver Figura 17.2, no *Capítulo 17*). O acréscimo de grupos metila às histonas (metilação da histona) também modifica a estrutura da cromatina, mas o efeito varia dependendo do aminoácido específico que é metilado, alguns tipos de metilação de histona estão associados a mais transcrição e outros tipos estão associados a menos transcrição. Por exemplo, a adição de três grupos metila à lisina 4 na histona H3 (H3K4me3, K representa lisina) é encontrada próxima de genes ativos na transcrição. A metilação da lisina 36 na histona H3 (H3K36me3) também está associada a mais transcrição. Por outro lado, a adição de três grupos metila à lisina 9 na H3 (H3K9me3) e à lisina 20 na histona 4 (H4K20me3) está associada à repressão da transcrição. Foi demonstrado que muitas marcas adicionais das histonas estão associadas ao nível de transcrição. Esses tipos de modificações são chamados de **marcadores epigenéticos**.

As modificações de histona são adicionadas e removidas por proteínas especiais. As proteínas do grupo polycomb (PcG) são um grande grupo de proteínas que reprimem a transcrição ao modificar as histonas. Essas modificações alteram a estrutura da cromatina de modo que o DNA não fica acessível para os fatores de transcrição, a RNA polimerase e outras proteínas necessárias para transcrição. Por exemplo, o complexo polycomb 2 repressivo (PRC2) adiciona dois ou três grupos metila à lisina 27 da histona H3, criando o marcador epigenético H3K27me3 que reprime a transcrição.

Muitas das enzimas e proteínas que produzem marcadores epigenéticos não podem se ligar a sequências específicas de DNA por si sós. Assim, precisam ser recrutadas para alvos específicos no cromossomo. Os fatores de transcrição específicos de sequência, modificações de histona preexistente e moléculas de RNA servem para recrutar enzimas modificadoras de histona para sítios específicos.

Pesquisa indica que as modificações de histona única, como as mencionadas aqui, não determinam individualmente a atividade de transcrição de um gene. Pelo contrário, é a presença combinada de múltiplos marcadores epigenéticos que determina o nível de atividade. Também existem consideráveis eventos de interferência (*crosstalk*) entre os marcadores epigenéticos, um marcador de histona pode afetar se marcadores adicionais ocorrem próximo e como eles funcionam. *Crosstalk* pode ocorrer porque as modificações de histona atraem enzimas e proteínas que modificam outras histonas. As modificações de histona não afetam apenas a transcrição, mas também podem influenciar outros processos moleculares como o reparo do DNA e a sinalização do ponto de verificação do ciclo celular (ver *Capítulo 23*). Por exemplo, a ubiquitinação da histona H2B é necessária para o reparo de quebras de fita dupla no DNA. Essa modificação leva a outras modificações de histona, como a metilação da lisina 79 de H3 (H3K79me3); essas modificações alteram a estrutura da cromatina e permitem acessar proteínas que reparam quebras de fita dupla.

Manutenção das modificações da histona. O processo pelo qual as modificações da histona são mantidas durante a divisão celular ainda não está bem compreendido como a metilação do DNA. Não existe mecanismo universal para manter as modificações da histona; diferentes tipos de modificações são indubitavelmente mantidos por mecanismos diferentes.

Foram propostos vários modelos para explicar como as modificações da histona são fielmente transmitidas para as células-filhas. Durante o processo de replicação do DNA, os nucleossomos são rompidos e as proteínas histona originais são distribuídas aleatoriamente entre as duas novas moléculas de DNA. As histonas recém-sintetizadas são adicionadas para completar a formação de novos nucleossomos (ver *Capítulo 12*). A maioria dos modelos assume que após a replicação os marcadores epigenéticos permanecem nas histonas originais e estas marcas recrutam enzimas que fazem mudanças semelhantes nas novas histonas. Por exemplo, PRC2 adiciona o marcador epigenético H3K27me3 nas histonas. PRC2 tem preferencialmente como alvo histonas na cromatina que já tenha um marcador H3K27me3, garantindo que qualquer novo nucleossomo que seja adicionado após a replicação também seja metilado. Dessa forma, as modificações da histona podem ser mantidas durante a divisão celular.

Conceitos

A modificação das proteínas histona, incluindo a adição de grupos metila, grupos acetila, fosfatos e a ubiquitinação alteram a estrutura da cromatina. Algumas dessas modificações são transmitidas para as células-filhas durante a divisão celular e para gerações futuras.

Efeitos epigenéticos produzidos pelas moléculas de RNA

Novos indícios demonstram que as moléculas de RNA são importantes para produzir os efeitos epigenéticos. O primeiro descoberto e ainda mais bem compreendido exemplo de mudança epigenética mediada por RNA é a inativação do X, na qual um longo RNA não codificado chamado *Xist* suprime a transcrição de um dos cromossomos X nas fêmeas dos mamíferos. Outro exemplo envolve a paramutação no milho, no qual um alelo epigeneticamente alterado induz uma mudança em outro alelo que então é transmitido para gerações futuras. A paramutação no milho é produzida por siRNAs (ver *Capítulo 14*). Ambos os exemplos serão discutidos posteriormente neste capítulo.

Diferentes mecanismos estão envolvidos nas mudanças epigenéticas pelas moléculas de RNA. No caso da inativação do X, o RNA *Xist* recobre um cromossomo X e então atrai a PRC2, que deposita grupos metila na lisina 27 da histona H3, criando o marcador epigenético H3K27me3 que altera a estrutura da cromatina e reprime a transcrição.

Ocorrem outros exemplos de fenótipos epigenéticos associados a RNA pelas moléculas siRNAs que silenciam genes e elementos de transposição (ver *Capítulo 18*) ao direcionar a metilação do DNA ou modificações da histona em sequências específicas de DNA. Além disso, a pesquisa demonstrou que os processos epigenéticos como metilação e modificação de histona influenciam a expressão dos microRNAs (ver *Capítulo 14*) que, por sua vez, são importantes para regular outros genes. MicroRNAs também controlam a expressão de genes que produzem efeitos epigenéticos, como as enzimas que metilam DNA e modificam as proteínas histona. Não está claro como as mudanças epigenéticas baseadas em RNA são mantidas pelas divisões celulares, embora aparentemente algumas envolvam pequenos RNAs que são transmitidos pelo citoplasma.

Conceitos

As moléculas de RNA produzem modificação da cromatina por meio de vários processos.

Checagem dos conceitos 2

Qual das opções não é um mecanismo principal da mudança epigenética?

a. Metilação do DNA.
b. A alteração de uma sequência de base de DNA em um promotor.
c. Acetilação de histona.
d. Ação das moléculas de RNA.

21.3 Os processos epigenéticos produzem um conjunto variado de efeitos

Inicialmente, acreditava-se que os mecanismos epigenéticos fossem importantes em um pequeno número de fenótipos incomuns. Entretanto, a pesquisa nos últimos 15 anos revelou que a epigenética sustenta um arranjo impressionante e até crescente de fenômenos biológicos. Vamos examinar alguns exemplos desses efeitos epigenéticos nesta seção.

Paramutação

Um dos primeiros exemplos de epigenética foi um fenótipo curioso que Alexander Brink descreveu no milho nos anos 1950. Brink estava estudando o *locus r1*, que ajuda a determinar a pigmentação nas sementes do milho. O alelo R^r nesse *locus* produz grãos púrpura, enquanto o alelo R^{st} codifica grãos manchados. Brink observou que quando R^{st} estava presente em um genótipo com o alelo R^r, o alelo R^{st} alterava de forma permanente a expressão do alelo R^r, de modo que agora ele também produzia sementes manchadas. Esse efeito reduzido do alelo R^r alterado sobre a pigmentação persistiu por várias gerações, mesmo na ausência do alelo R^{st}. Brink chamou esse fenômeno de paramutação. A paramutação viola o princípio de segregação de Mendel, que afirma que quando os gametas são formados, cada alelo se separa e é transmitido de forma independente para a próxima geração.

Atualmente, a **paramutação** é definida como uma interação de dois alelos que leva a uma mudança que pode ser herdada na expressão de um dos alelos. É surpreendente que a paramutação produza essas diferenças no fenótipo sem qualquer alteração na sequência de bases do DNA do alelo convertido. O fenômeno da paramutação tem várias características importantes. Primeiro, o padrão de expressão recém-estabelecido do alelo convertido é transmitido para as gerações futuras, mesmo que o alelo que produza a alteração não esteja mais presente. Segundo, o alelo alterado agora é capaz de converter outros alelos para o novo fenótipo. E terceiro, não existem diferenças associadas de sequência de DNA nos alelos alterados. Foram descobertos vários exemplos de paramutação em diferentes organismos, e os geneticistas começaram a solucionar o mecanismo molecular desse curioso fenômeno.

Paramutação no milho. Alguns anos após Brink relatar a paramutação no *locus r1* do milho, outro exemplo relacionado foi

descoberto por Ed Coe Jr. Esse caso envolveu a interação de alelos no *locus b1* do milho, que também ajuda a determinar a pigmentação. A paramutação no *locus b1* é mais direta do que no *locus r1*, então usaremos o *b1* para examinar o processo e o mecanismo da paramutação.

O *locus b1* ajuda a determinar a quantidade de antocianina púrpura que o milho produz. O *locus* codifica um fator de transcrição que regula os genes envolvidos na produção do pigmento. As plantas homozigotas para o alelo *B-I* (*B-I B-I*) mostram elevada expressão do *locus b1* e são púrpura-escuras (**Figura 21.4**). As plantas homozigotas para o alelo *B′* (*B′B′*) têm uma expressão menor de *b1* e são levemente pigmentadas. Entretanto, as sequências de DNA dos alelos *B-I e B′* são idênticas. Os alelos geneticamente idênticos como estes, que produzem diferenças hereditárias nos fenótipos pelos processos epigenéticos, são chamados de **epialelos**.

Nas plantas heterozigotas *B-I B′*, o alelo *B-I* é convertido em *B′*, com o resultado de que as plantas heterozigotas são levemente pigmentadas (ver Figura 21.4), como as homozigotas *B′B′*. O alelo recém-convertido é indicado como *B′**. E importante, não existe diferença funcional entre *B′* e *B′**; o alelo *B′** agora é totalmente capaz de converter outros alelos *B-I* em alelos *B′** nas gerações seguintes.

Pesquisa demonstrou que uma das características necessárias para a paramutação no *locus b1* é a presença de uma série de sete sequências repetidas em *tandem* que estão aproximadamente 100.000 pares de bases *upstream* da sequência codificadora para o *locus b1* (**Figura 21.5**). Cada repetição tem 853 pb e não codifica nenhuma proteína. Ambos os alelos *B-I* e *B′* têm essas repetições em *tandem*, mas a estrutura da cromatina dos dois alelos é diferente: o alelo *B-I* tem a cromatina mais aberta. As repetições em *tandem* são necessárias para alta expressão do alelo *B-I* e a produção de pigmento. Foi sugerido que as repetições atuam como um acentuador (ver *Capítulo 17*), estimulando a transcrição no *locus b1*, mas apenas quando a cromatina que circunda as repetições está em uma configuração aberta, como está no alelo *B-I*. A configuração mais fechada do alelo *B′* pode evitar as repetições de interagir com o promotor de *b1* e estimular a transcrição. Não se sabe como as repetições podem interagir com o alelo *B′*.

Os diferentes estados da cromatina de *B-I* e *B′* podem explicar seus diferentes níveis de expressão, mas como o alelo B converte o alelo *B-I* em *B′**? Embora o mecanismo não esteja totalmente compreendido, pesquisa recente demonstra que é provável que a comunicação entre *B′* e *B-I* ocorra pela ação de pequenas moléculas de RNA.

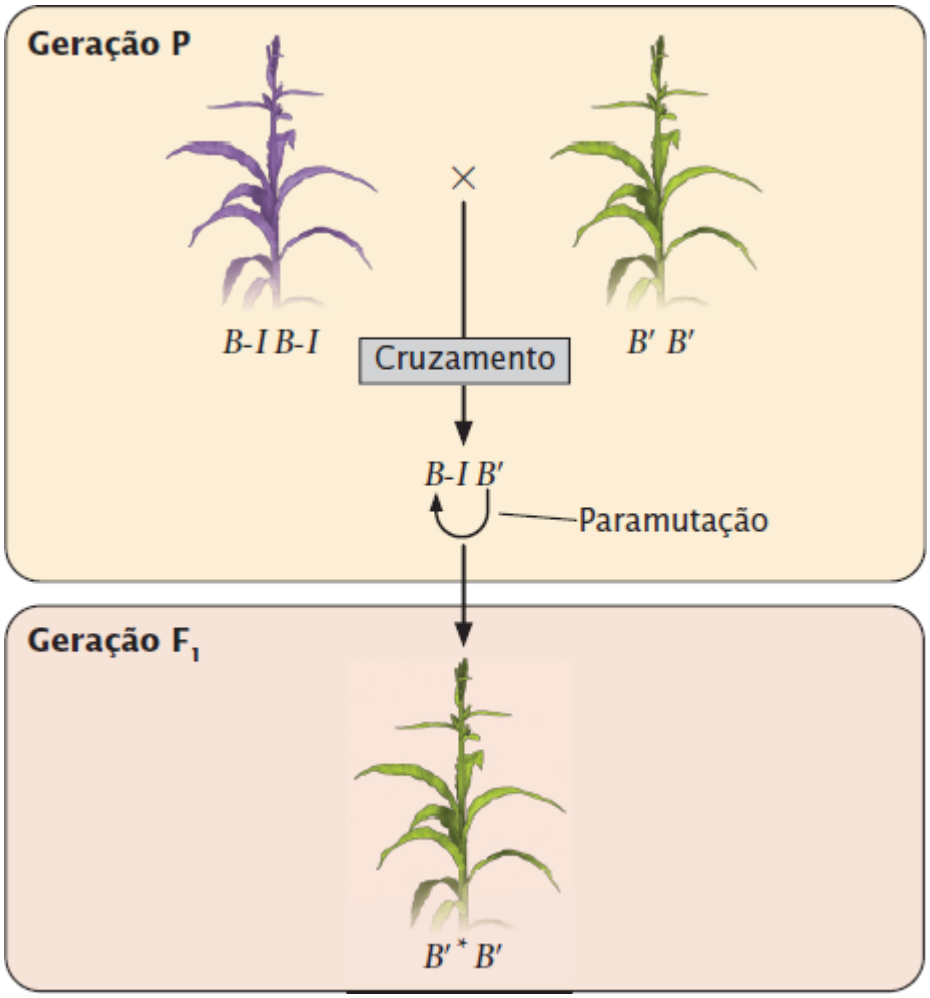


Figura 21.4 Na paramutação no *locus b1* do milho, uma cópia do alelo *B′* converte o alelo *B-I em B′, que tem o mesmo fenótipo que B′.** O genótipo *B-I B-I* produz uma planta pigmentada, enquanto as plantas com os genótipos *B′B′* e *B′B′** são levemente

pigmentadas.

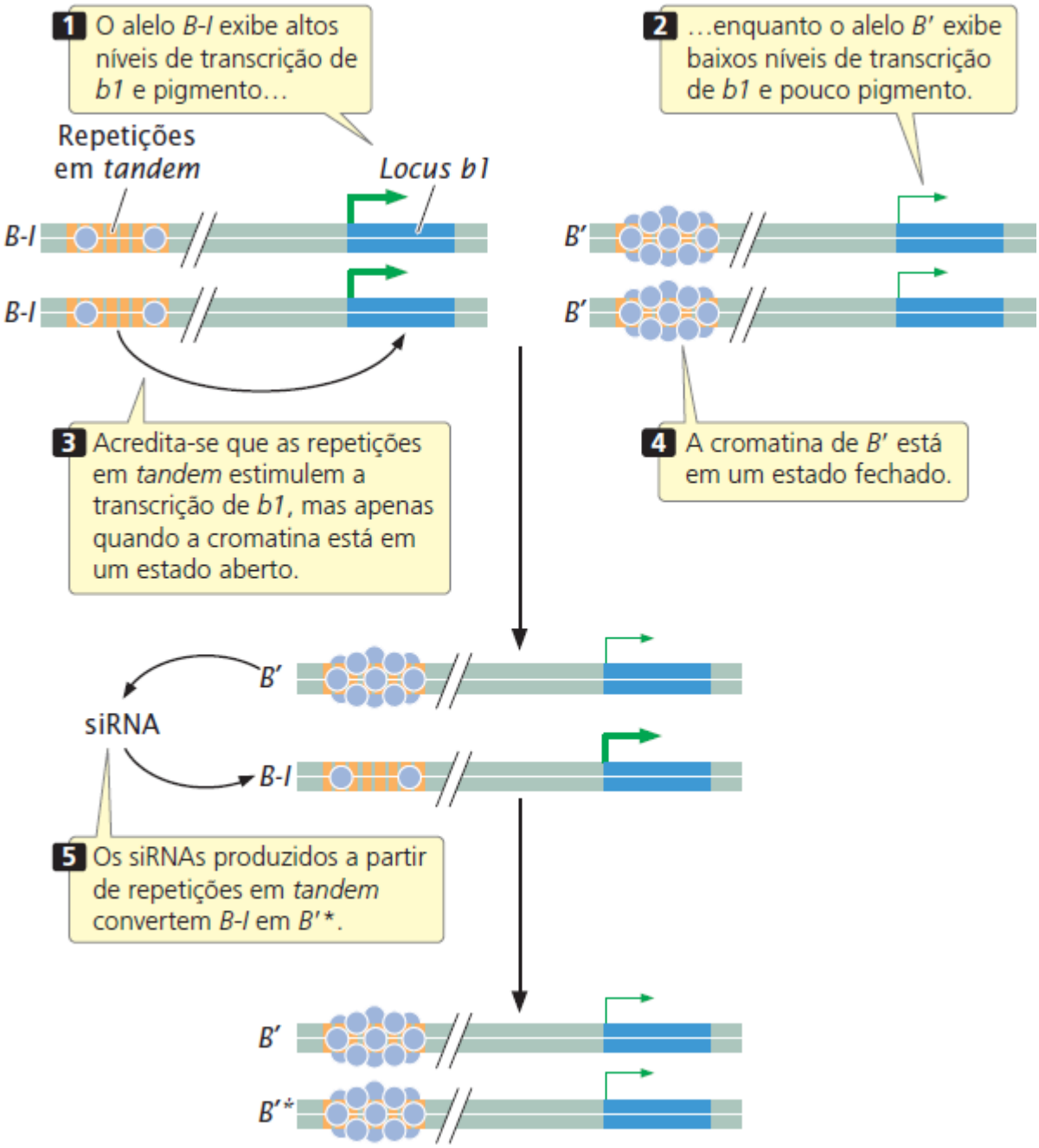


Figura 21.5 A paramutação no *locus b1* no milho requer a presença de 7 repetições em *tandem upstream* deste *locus*.

As repetições em *tandem* necessárias para a paramutação codificam siRNAs com 25 nucleotídios (ver *Capítulo 14*). Alguns siRNAs são conhecidos por modificar a estrutura da cromatina ao direcionar a metilação do DNA a sequências específicas de DNA. Os geneticistas isolaram vários genes no milho necessários para a que ocorra a paramutação; inativar esses genes elimina o fenômeno. Um desses genes é *mop1*, que codifica uma RNA polimerase direcionada para o RNA (uma enzima que sintetiza RNA a partir de um molde de RNA). Esse gene é necessário para gerar os siRNAs codificados pelas repetições em *tandem*, embora não pareça que essa enzima transcreva de fato as cópias de DNA das repetições em *tandem*. Outro gene necessário para a paramutação, chamado *rmr1*, codifica uma proteína remodeladora de cromatina.

Dessa forma, evidências atuais sugerem que as moléculas de siRNA convertem *B-I* em *B′**, e é provável que essa conversão envolva uma mudança nos estados de cromatina dos alelos. Existem outros exemplos de plantas em que os siRNAs influenciam a metilação do DNA e a estrutura da cromatina. Entretanto, não se sabe exatamente como as moléculas de siRNAs produzem essa mudança na cromatina durante a paramutação. É provável que o estado alterado da cromatina das repetições no alelo *B′* reduza a transcrição do *locus b1*, talvez ao interferir com a interação das repetições em *tandem* e o promotor de *b1*. Supõe-se que a alteração na cromatina do alelo *B′* seja estável e transmitida para gerações futuras. A pesquisa mostra que a transcrição das repetições em *tandem* e a geração de siRNAs a partir destas são necessárias, mas não suficientes para a paramutação, então fatores adicionais devem estar envolvidos. Também não está claro como a produção de siRNAs é transmitida entre as gerações.

Paramutação nos camundongos. Também foram observados vários exemplos de paramutação nos camundongos. Uma envolve o *locus Kit*, que codifica um receptor da tirosina quinase e atua na produção de pigmento, desenvolvimento de célula germinativa e produção de células vermelhas. Os geneticistas criaram, por engenharia genética, um alelo *Kit* mutante (chamado aqui de Kit^t), que carreia uma porção de 3.000 pb do gene *lacZ* (ver *Capítulo 16*) inserido no *locus Kit*. Os camundongos que são homozigotos para o alelo do tipo selvagem (Kit^+ Kit^+) têm pigmento normal. Os camundongos que são homozigotos para os alelos *Kit* mutantes (Kit^t Kit^t) morrem logo após nascer. Os camundongos heterozigotos para os alelos do tipo selvagem e mutante (Kit^+ Kit^t) têm pontas de cauda e patas brancas (**Figura 21.6**). Quando um camundongo heterozigoto é cruzado com um camundongo homozigoto do tipo selvagem, metade dos descendentes é homozigota (Kit^+ Kit^+) e metade é heterozigota (Kit^+ Kit^t), como seria esperado. Entretanto, muitos dos camundongos Kit^+ Kit^+ desenvolvem patas e caudas brancas, o fenótipo esperado dos heterozigotos. Na presença do alelo Kit^t no heterozigoto, o alelo Kit^+ é alterado de forma que ele tem o mesmo fenótipo de Kit^t. Os camundongos com esses alelos alterados são indicados como *Kit**. O alelo *Kit** alterado é transmitido de forma estável para as gerações futuras, nas quais continua a produzir caudas e patas brancas. Alguns humanos com uma mancha branca na testa e áreas de pigmento reduzido (chamado de traço de cavalo malhado) têm mutações no *locus Kit*; outras mutações no *Kit* resultam em predisposição para alguns cânceres.

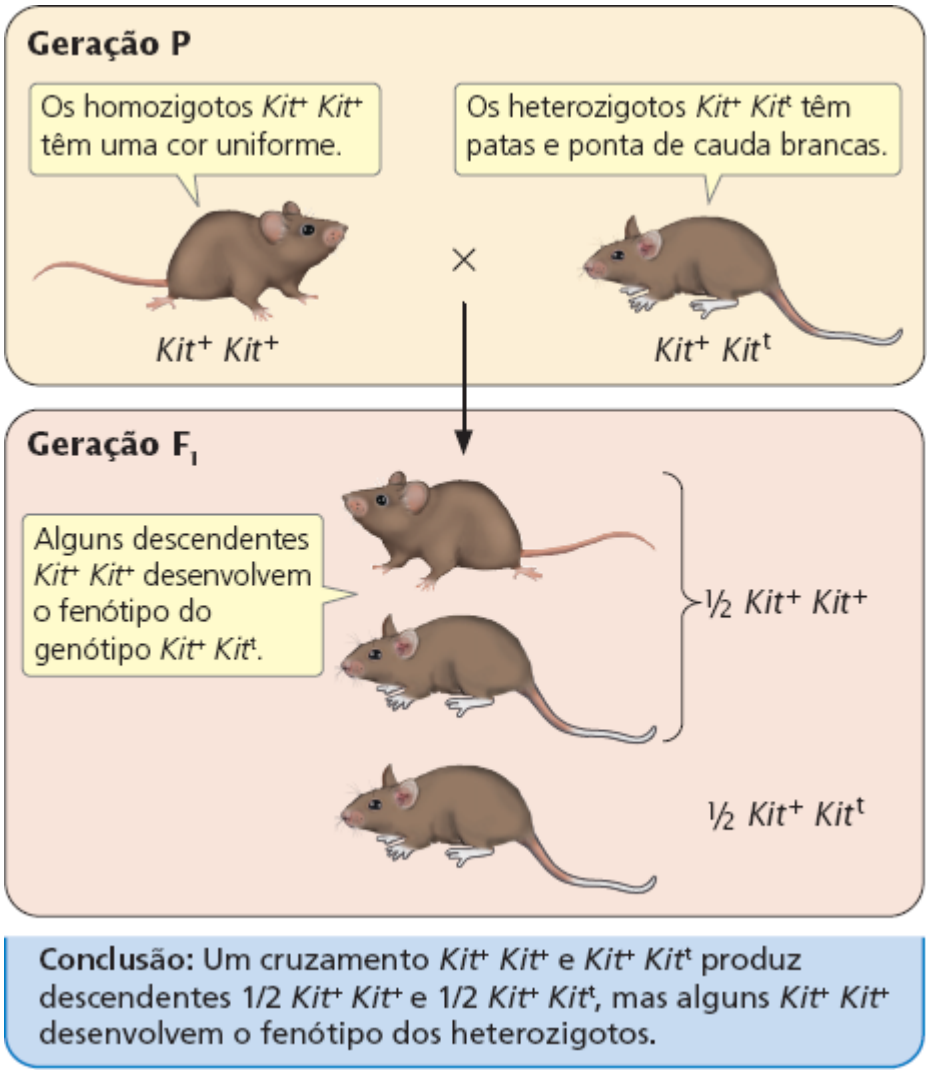


Figura 21.6 A paramutação no *locus Kit* dos camundongos.

Pesquisadores demonstraram que esse exemplo de paramutação também é mediado por molécula de RNA, embora seja provável que o mecanismo seja diferente do observado na paramutação no milho. Nos camundongos, as patas e caudas brancas do alelo Kit^t parece ser causado por microRNAs (miRNAs) que degradam o mRNA *Kit* e esses miRNAs são transmitidos para as gerações futuras pelos gametas. Pesquisadores observaram uma redução pela metade no mRNA *Kit* nos camundongos heterozigotos e camundongos *Kit**, sugerindo que as caudas e patas brancas dos heterozigotos sejam o resultado da redução na quantidade de mRNA *Kit*. Para determinar se o RNA era responsável pela paramutação, eles injetaram em alguns embriões do tipo selvagem o RNA de homozigotos Kit^+ Kit^+ e injetaram em outros embriões do tipo selvagem o RNA de heterozigotos Kit^+ Kit^t. Entre os camundongos que tiveram o desenvolvimento completo, eles observaram que as caudas e patas brancas eram mais frequentes nos que receberam a injeção de RNAs dos heterozigotos, sugerindo que o RNA dos heterozigotos é capaz de alterar

o alelo *Kit*$^{+}$ dos camundongos do tipo selvagem (**Quadro 21.1**). Os pesquisadores injetaram miRNAs que degradam o mRNA *Kit* em embriões do tipo selvagem. Esse experimento produziu mais camundongos com patas e caudas brancas do que quando injetaram miRNAs inespecíficos em embriões (ver Quadro 21.1). A capacidade de produzir caudas e patas brancas características do genótipo *Kit*$^{+}$ *Kit*^{t} ao injetar miRNA em embriões sugere que esse caso de paramutação está associado a moléculas de miRNA que são transferidas para o embrião pelo óvulo e espermatozoide. Entretanto, ainda existem muitos aspectos desconhecidos da paramutação no *locus Kit*.

Conceitos

A paramutação ocorre quando um alelo cria uma alteração hereditária em outro alelo sem qualquer mudança na sequência de DNA. Pesquisa sugere que a paramutação no milho e camundongos é mediada por pequenas moléculas de RNA.

Checagem dos conceitos 3

Qual dos itens a seguir é uma característica da paramutação?

a. Um alelo é capaz de alterar outro alelo quando ambos estão presentes em um heterozigoto.
b. Os alelos alterados podem ser transmitidos para as gerações futuras.
c. Os alelos alterados devem ser capazes de alterar outros alelos em futuras gerações.
d. Todas as opções anteriores.

Quadro 21.1 Efeitos da injeção de diferentes tipos de RNA em camundongos do tipo selvagem.*

Tipo de RNA injetado	Presença de pontas de cauda e patas brancas
mRNA *Kit*$^{+}$ *Kit*$^{+}$	Incomum
mRNA *Kit*$^{+}$ *Kit*^{t}	Mais comum
miRNA para mRNA *Kit*	Mais comum
miRNA inespecífico	Incomum

**Kit*$^{+}$ *Kit*$^{+}$

Epigenética comportamental

A pesquisa mostrou que as experiências de vida, especialmente as obtidas no início da vida, podem ter efeitos duradouros no comportamento e, em alguns casos, nas gerações futuras. Os pesquisadores estão descobrindo que esses efeitos duradouros são mediados por processos epigenéticos. O número de estudos que demonstram de forma convincente que a experiência de vida altera a estrutura da cromatina é pequeno no momento (e alguns ainda são controversos), mas vários pesquisadores estão procurando por efeitos epigenéticos de experiência e seus efeitos duradouros na estrutura da cromatina e no comportamento.

Mudanças epigenéticas induzidas pelo comportamento materno. Um exemplo fascinante de epigenética comportamental é observado nos efeitos duradouros do comportamento maternal nos ratos. Uma fêmea lambe e cuida de seus descendentes, enquanto ela se curva e os amamenta. Os descendentes das fêmeas que apresentam um comportamento mais evidente de lamber e cuidar têm menos medo quando adultos e mostram respostas hormonais reduzidas ao estresse comparados com os descendentes de fêmeas que lambem e cuidam menos. Essas diferenças duradouras nos descendentes não são devido a diferenças genéticas herdadas de suas genitoras, pelo menos não as diferenças genéticas nas sequências de bases do DNA. Os descendentes expostos a mais lambidas e cuidados desenvolvem um padrão diferente de metilação do DNA em comparação com os descendentes com menos lambidas e cuidados. Essas diferenças na metilação do DNA afetam a acetilação das proteínas histona que persistem na fase adulta e alteram a expressão do gene do receptor de glicocorticoide, que é importante nas respostas hormonais ao estresse. A expressão de outros genes de resposta ao estresse também é afetada.

Para demonstrar o efeito da estrutura da cromatina alterada à resposta de estresse dos descendentes, os pesquisadores introduziram um inibidor de desacetilase nos cérebros dos ratos jovens, que evita a remoção dos grupos acetila das proteínas histona. Após a infusão desse inibidor, as diferenças na metilação do DNA e acetilação da histona associadas ao comportamento de cuidados desapareceram, assim como a diferença nas respostas a medo e estresse nos adultos. Isto demonstra que o comportamento de lamber e cuidar do camundongo fêmea produz mudanças epigenéticas na cromatina dos descendentes, que

provoca diferenças duradouras no seu comportamento. Resolva o Problema 28

Efeitos epigenéticos do estresse precoce nos humanos. Numerosos estudos demonstraram que o estresse durante a infância e a adolescência provoca vários efeitos adversos que perduram na vida adulta. Por exemplo, abuso na infância aumenta a probabilidade de que a criança experimentará depressão, ansiedade e suicídio quando for um adulto. Em um estudo, os pesquisadores examinaram os cérebros de 24 pessoas que cometeram suicídio, metade das quais haviam sofrido maustratos na infância. Eles descobriram que as que sofreram maus-tratos na infância tinham um grau maior de metilação do gene do receptor do glicocorticoide, um gene envolvido na resposta ao estresse, do que as que não sofreram maus-tratos. Embora o número de cérebros estudados fosse pequeno, o estudo sugere que o estresse na primeira infância pode de fato provocar modificações epigenéticas na estrutura da cromatina nos humanos.

Outros estudos demonstraram que a expressão gênica é afetada por experiência de vida precoce. Por exemplo, pesquisadores descobriram que crescer em um ambiente socioeconômico inferior antes dos 5 anos alterou a expressão de mais de 100 genes relacionados com a função imunológica dos adultos. A introdução do Capítulo 11 discute a observação de que o estresse na primeira infância, como crescer em um orfanato, altera o comprimento do telômero, um tipo de mudança epigenética.

Epigenética na cognição. Vários estudos de pesquisa demonstraram que anormalidades na metilação do DNA estão associadas a distúrbios de desenvolvimento e capacidade intelectual nos humanos. Esses achados levaram os pesquisadores a procurar por efeitos da estrutura da cromatina sobre o aprendizado, a memória e a capacidade cognitiva nos camundongos e ratos. Um estudo descobriu que treinar o camundongo a evitar estímulo adverso em uma localização específica reduzia metilação do DNA do gene *Bdnf*, que codifica um fator de crescimento que estimula o crescimento de conexões entre os neurônios. Quando desmetilado, o gene *Bndf* era mais ativo. Quando os pesquisadores injetaram nos cérebros dos camundongos um fármaco que inibia a desmetilação, a atividade do gene *Bdnf* era reduzida, e a memória dos camundongos sobre onde o estímulo adverso ocorria também reduziu.

Outro estudo descobriu que um fármaco que promovia a acetilação das proteínas histona melhorou o aprendizado e memória nos camundongos que tinham um distúrbio semelhante à doença de Alzheimer. A acetilação das histonas altera a estrutura da cromatina ao afrouxar a associação do DNA com as proteínas histona e estimular a transcrição de vários genes. Outros estudos descobriram que a acetilação da histona reduz com a idade nos camundongos, com expressão diminuída de genes relacionados com o aprendizado e a memória. Quando os pesquisadores injetaram em camundongos um fármaco inibidor da atividade da desacetilase, a acetilação das histonas aumentou, a transcrição dos genes envolvidos com a memória aumentou e a memória dos camundongos melhorou. Esses estudos sugerem que mudanças na estrutura da cromatina podem estar envolvidas com a memória e o aprendizado.

Conceitos

Alguns estudos estão fornecendo indícios de que as experiências no início da vida podem produzir mudanças epigenéticas que têm efeitos duradouros sobre o comportamento.

Efeitos epigenéticos das substâncias químicas do ambiente

Como algumas substâncias químicas são capazes de modificar a estrutura da cromatina, os pesquisadores procuram por efeitos duradouros das substâncias tóxicas do ambiente sobre a estrutura da cromatina e os traços epigenéticos.

Existe muito interesse recente em substâncias químicas, chamadas desreguladores endócrinos, que mimetizam ou interferem com os hormônios naturais. Os desreguladores endócrinos são capazes de interferir com processos regulados por hormônios naturais, como o desenvolvimento sexual e reprodução. Um desregulador endócrino é a vinclozolina, um fungicida comum usado no controle de fungos nas uvas, frutas e vegetais e para tratar a grama de campos de golfe. A vinclozolina atua como um antagonista no receptor de andrógeno – essa substância e seus metabólitos mimetizam a testosterona e se ligam ao receptor de andrógeno, evitando que a testosterona se ligue. Mas a vinclozolina e seus metabólitos não ativam adequadamente o receptor e dessa forma, ela inibe a ação dos andrógenos e impede a produção de espermatozoides.

Em um estudo, os pesquisadores descobriram que a exposição de ratos embrionários à vinclozolina levou a produção reduzida de espermatozoides não apenas nos animais tratados (quando eles alcançavam a puberdade), mas também em várias gerações subsequentes. Foi observada metilação aumentada do DNA nos espermatozoides dos machos expostos à vinclozolina e esses padrões de metilação foram herdados. Esse estudo e outros levantaram questões que, por meio das mudanças epigenéticas, a exposição ambiental a algumas substâncias químicas poderia ter efeitos na saúde das gerações futuras. Resolva o Problema 29

Conceitos

Por meio das mudanças epigenéticas, as substâncias químicas ambientais podem ter influências que se estendem para gerações posteriores.

Efeitos epigenéticos transgeração sobre o metabolismo

Na introdução deste capítulo, discutimos como a dieta durante a infância pode ter efeitos na saúde que podem transmitidos para as futuras gerações. Esses tipos de estudos epidemiológicos nos humanos são apoiados por estudos laboratoriais de camundongos e ratos. Em um estudo, os pesquisadores alimentaram camundongos machos *inbred* (isogênicos, consanguíneos) com dieta normal (controle) ou hipoproteica. Eles cruzaram os machos de ambos os grupos com fêmeas do controle alimentadas com dieta normal. Os machos foram então separados das fêmeas e nunca mais tiveram contato com seus descendentes; sua única contribuição para os descendentes foi um conjunto de genes transferidos pelos espermatozoides.

Os descendentes foram criados e seus níveis sanguíneos de lipídios e colesterol foram examinados. Os descendentes dos machos que receberam dieta hipoproteica exibiam expressão aumentada de genes envolvidos no metabolismo de lipídios e colesterol e correspondente redução dos níveis de colesterol, comparados com os descendentes dos machos que receberam dieta normal. Eles também observaram numerosas diferenças na metilação do DNA nos descendentes dos dois tipos de genitores, embora não pudessem ser encontradas diferenças nos padrões de metilação dos espermatozoides dos dois grupos de genitores. Esses resultados sugerem que as mudanças epigenéticas alteraram o metabolismo de colesterol dos descendentes, embora não esteja claro como as diferenças na metilação foram transmitidas dos genitores para os descendentes.

Em outro estudo, os pesquisadores alimentaram ratos machos com uma dieta hiperlipídica e, obviamente, eles ganharam peso. Esses machos foram cruzados com fêmeas que receberam dieta normal. Os descendentes também receberam dieta normal. As descendentes dos ratos machos que receberam dieta hiperlipídica tinham peso normal, mas na vida adulta desenvolveram intolerância à glicose e comprometimento da secreção de insulina (semelhante ao diabetes melito). Os pesquisadores observaram que nas células das ilhotas pancreáticas secretoras de insulina das descendentes, a expressão dos 642 genes envolvidos na secreção de insulina e tolerância à glicose foi alterada, demonstrando que a dieta do genitor afetou a expressão gênica em suas descendentes.

Efeitos epigenéticos em gêmeos monozigóticos

Os gêmeos monozigóticos (idênticos) se desenvolvem a partir de um único óvulo fertilizado por um espermatozoide que se divide e dá origem a dois zigotos (ver *Capítulo 6*). Os gêmeos monozigóticos são geneticamente idênticos, no sentido de que têm sequências de DNA idênticas, mas eles apresentam diferenças na aparência, na saúde e no comportamento. A natureza dessas diferenças nos fenótipos dos gêmeos idênticos ainda não está bem compreendida, mas evidências recentes sugerem que pelo menos algumas dessas diferenças são consequentes a alterações epigenéticas. Em um estudo, Mario Fraga, no Spanish National Cancer Center, e seus colaboradores examinaram 80 gêmeos idênticos e compararam o grau e a localização de metilação do seu DNA e acetilação de histona. Eles descobriram que a metilação de DNA e a acetilação de histona nos gêmeos idênticos eram semelhantes no início da vida, mas os gêmeos mais velhos tinham diferenças marcantes no conteúdo total e na distribuição de metilação de DNA e acetilação de histona. Além disso, essas diferenças afetavam a expressão gênica nos gêmeos. Essa pesquisa sugere que os gêmeos idênticos são epigeneticamente diferentes e que as diferenças fenotípicas entre eles podem ser causadas por expressão gênica diferenciada.

Conceitos

As diferenças fenotípicas entre gêmeos monozigóticos geneticamente idênticos resultam de efeitos epigenéticos.

Checagem dos conceitos 4

Que grau de diferença você esperaria observar nas sequências de bases de DNA e nos marcadores epigenéticos dos gêmeos monozigóticos?

a. Diferenças semelhantes na sequência de bases de DNA e nos marcadores epigenéticos.
b. Diferenças maiores na sequência de bases de DNA do que nos marcadores epigenéticos.
c. Diferenças maiores nos marcadores epigenéticos do que na sequência de bases de DNA.
d. Nenhuma diferença na sequência de bases de DNA ou nos marcadores epigenéticos.

Inativação do X

Nas fêmeas dos mamíferos, um cromossomo X em cada célula é aleatoriamente inativado para fornecer uma expressão igual dos genes ligados ao X nos machos e fêmeas (ver *Capítulo 4*). Por esse processo, chamado de inativação do X, muitos genes no cromossomo X inativado são silenciados de forma permanente e não são transcritos. Depois que um cromossomo X é inativado em uma célula, ele permanece inativado quando o DNA é replicado, e a marca da inativação é transmitida para as células-filhas durante a mitose. Esse fenômeno é responsável pela distribuição desigual de pigmento preto e laranja observada nos gatos com pelagem tricolor ou casco de tartaruga (ver *Capítulo 4*). A inativação do X é um tipo de efeito epigenético porque ele resulta em uma mudança estável na expressão gênica que é transmitida para as outras células.

Grande parte da pesquisa demonstrou que qual cromossomo X é inativado dentro de uma célula é controlado por um segmento específico do cromossomo X chamado de **centro de inativação do X**, que tem 100.000 a 500.000 pb de comprimento. A inativação é iniciada nesse centro e então se espalha pelo restante do cromossomo X inativado. O exame do centro de inativação do X levou à descoberta de vários genes importantes para inativar um cromossomo X em cada célula da fêmea e manter o outro cromossomo X ativo (**Figura 21.7**).

O ator principal na inativação do X é um gene chamado *Xist* (transcrito específico de inativação do X) que codifica um longo RNA não codificador (lncRNA) com 17.000 pb de comprimento (**Figura 21.8**). Como seu nome sugere, essa molécula de RNA não codifica proteína. Pelo contrário, o lncRNA do *Xist* recobre o cromossomo X de onde ele foi transcrito. O lncRNA *Xist* então atrai o complexo repressor polycomb 2 (PRC2) e, consequentemente, o complexo repressor polycomb 1 (PRC1). Essas proteínas produzem marcadores epigenéticos, como a trimetilação da lisina 27 da histona 3 (H3K27me3), e outras modificações de histona que reprimem a transcrição. Consequentemente, muitos dinucleotídios CpG são metilados, levando ao permanente silenciamento do cromossomo X inativado.

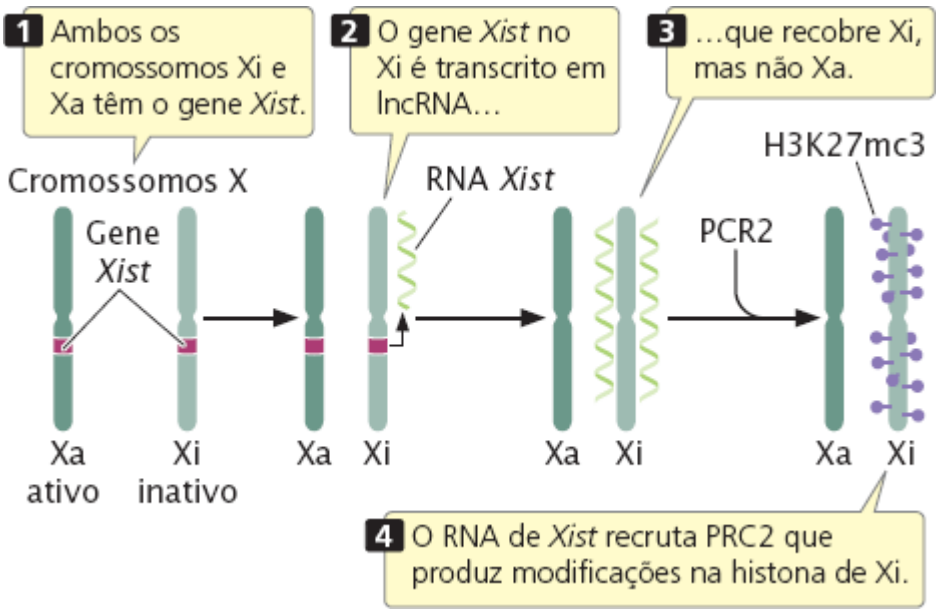


Figura 21.7 Na inativação do X, o gene *Xist* no X inativo produz um longo RNA não codificado que recobre o cromossomo X inativo e suprime a transcrição.

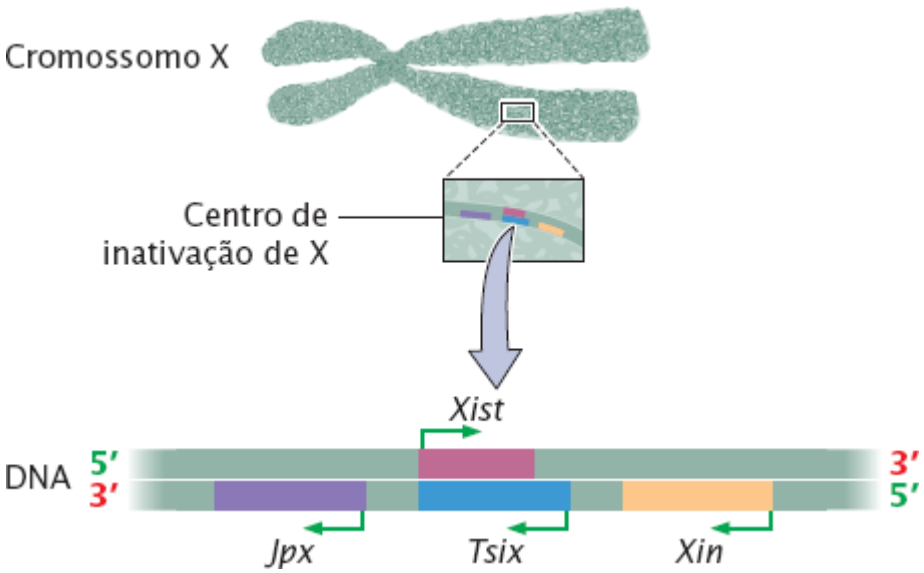


Figura 21.8 Vários genes no centro de inativação de X interagem para produzir a inativação de um cromossomo X enquanto mantêm o outro cromossomo X ativo.

Nos camundongos, eles são dois eventos de inativação separados. Logo após a fertilização, quando o embrião chega ao estágio de 8 células, o cromossomo X do genitor masculino é inativado, enquanto o cromossomo X materno permanece ativo. No embrião em desenvolvimento, o cromossomo X paterno é reativado durante a maturação do blastocisto. A inativação ocorre novamente no desenvolvimento inicial, mas agora qual X é inativado é aleatório – o X da genitora e o X do genitor têm a mesma probabilidade de serem inativados. A partir desse ponto, qualquer X inativado permanece silenciado durante as divisões celulares subsequentes. Entretanto, alguns genes do cromossomo X inativado escapam da inativação e continuam a ser transcritos. Como esses genes escapam da inativação do X é desconhecido. O mais interessante, nos mamíferos marsupiais, o cromossomo X paterno é a cópia que permanece silenciada de forma permanente em todas as células.

Como foi mencionado, a inativação do X é feita pela transcrição do gene *Xist* do cromossomo X inativo para produzir o lncRNA *Xist*, que recobre o cromossomo X inativo e permite as mudanças na estrutura da cromatina que silenciam a transcrição. Mas o que acontece no cromossomo X ativo? Por que ele não é recoberto pelo RNA do *Xist* e silenciado? Embora todos os detalhes dess processo ainda não estejam bem compreendidos, pesquisa recente demonstrou que existem vários genes adicionais no centro de inativação de X que codificam outros lncRNAs. Esses lncRNAs ajudam a produzir a inativação do X do X inativo, enquanto não silenciam o X ativo (ver Figura 21.7). Um destes é o gene *Tsix*, que é transcrito no cromossomo X ativo. *Tsix* é o antissenso de *Xist*, o que significa que ele se sobrepõe ao gene *Xist* e é transcrito a partir da fita oposta (ver Figura 21.8), produzindo o lncRNA *Tsix* que é complementar ao lncRNA *Xist*. Por meio de vários mecanismos, o *Tsix* reprime a expressão do *Xist* no cromossomo X ativo. Outro ator importante é um gene chamado *Jpx*, que codifica um lncRNA que estimula a transcrição de *Xist* no cromossomo X inativo. Assim, *Xist* é controlado por duas chaves paralelas com efeitos opostos: (1) *Jpx* estimula a expressão de *Xist* no cromossomo X inativo, fazendo com que *Xist* seja transcrito e leve à inativação de X e (2) *Tsix* reprime *Xist* no cromossomo X ativo, fazendo com que *Xist* não seja transcrito naquele cromossomo e evitando a inativação. Vários outros genes também estão envolvidos. Um gene chamado *Xite* codifica um lncRNA que mantém a expressão de *Tsix* no cromossomo X ativo. O **Quadro 21.2** resume os principais genes envolvidos no processo de inativação de X.

Esse processo complexo garante que um cromossomo X esteja inativo e um permaneça ativo em cada célula da fêmea. Os cientistas já reconhecem que a inativação do X também envolve algum tipo de mecanismo capaz de contar os cromossomos X, porque todos, exceto um cromossomo X em cada célula, estão inativados. Assim, o único X nas células de um macho XY permanece ativo (sem ocorrer ativação de X) e dois cromossomos são inativados nas fêmeas XXX (ver a discussão sobre corpúsculos de Barr no *Capítulo 4*). A natureza desse mecanismo de contagem ainda não está bem compreendida. **Resolva o Problema 31**

Conceitos

As mudanças epigenéticas sustentam a inativação de X, no qual um cromossomo X nas células das fêmeas é silenciado de forma permanente. A inativação de X ocorre pela ação de vários genes no centro de inativação de X que codificam longos RNAs não codificados. Os produtos desses genes interagem para garantir que um cromossomo X esteja inativado e um permaneça ativo em cada célula da fêmea.

Checagem dos conceitos 5

Qual seria o efeito de introduzir siRNAs que degradam o RNA *Xist* em uma célula de fêmea?

Quadro 21.2 Principais genes envolvidos na inativação de X.

Gene	Codifica	Ação do gene
Xist	lncRNA	Reveste o cromossomo X inativo e leva ao silenciamento da transcrição de vários genes no X inativo
Tsix	lncRNA	Inibe a transcrição de *Xist* no cromossomo X ativo

Jpx	lncRNA	Estimula a transcrição de *Xist* no cromossomo X inativo
Xite	lncRNA	Mantém a expressão de *Tsix* no X ativo, que inibe *Xist* e mantém a transcrição dos genes no cromossomo X ativo

Alterações epigenéticas associadas à diferenciação celular

Todas as células do corpo humano são geneticamente idênticas, ainda assim, diferentes tipos de células exibem incríveis fenótipos diferentes; uma célula nervosa é bem diferente em tamanho, formato e função de uma célula intestinal. Essas diferenças nos fenótipos são estáveis e transmitidas de uma célula para outra, apesar do fato de que as sequências de DNA de todas as células são as mesmas.

As **células-tronco** são células não diferenciadas capazes de formar todo tipo de célula em um organismo, uma propriedade chamada de **pluripotência**. À medida que as células-tronco se dividem e dão origem a um tipo de célula mais especializado, o programa de expressão gênica da célula se torna progressivamente fixo, de modo que cada tipo celular expressa apenas os genes necessários para fazer as funções daquele tipo de célula. Embora o controle desses programas de expressão específicos das células ainda não esteja bem compreendido, as mudanças na metilação do DNA e na estrutura de cromatina são importantes para silenciar alguns genes e ativar outros.

As células-tronco fornecem uma fonte potencial de células para regeneração de tecidos, tratamento médico e pesquisa. No passado, a única fonte de células-tronco com a capacidade de diferenciação nos tecidos adultos eram células de embriões, mas por causa de questões éticas sobre a criação e o uso de embriões humanos para cultivar células-tronco, os pesquisadores estão procurando há muito tempo a capacidade de induzir as células somáticas adultas em se desdiferenciar e reverter em célulastronco. Tais células são chamadas de **células-tronco pluripotentes induzidas** (iPSCs). Atualmente os pesquisadores conseguiram criar iPSCs graças ao tratamento de fibroblastos (células do tecido conjuntivo humano totalmente diferenciadas) em cultura com um coquetel de fatores de transcrição (**Figura 21.9**), embora menos de 1% das células tratadas reverta para iPSCs. Os fatores de transcrição que induzem a pluripotência provocam substancial reprogramação epigenética, alterando os padrões de metilação de DNA e modificações de histona que se acumulam com a diferenciação celular. Todavia, a pesquisa recente mostrou que as iPSCs retêm memória de seu passado e não são completamente equivalentes às células-tronco embrionárias (aquelas coletadas de embriões). Um estudo descobriu que, embora os padrões de metilação do DNA de iPSCs sejam muito diferentes dos padrões das células somáticas, as iPSCs conservam algumas marcas de metilação das células somáticas e que a metilação das iPSCs não era idêntica à metilação das células-tronco embrionárias. Outro estudo comparou as modificações de histona dos fibroblastos, iPSCs e células-tronco embrionárias. As iPSCs e células-tronco embrionárias tinham menos marcas H3K27me3 e H3K9me3 que os fibroblastos, mas os pesquisadores também descobriram mais dessas marcas importantes nas iPSCs do que nas células-tronco embrionárias.

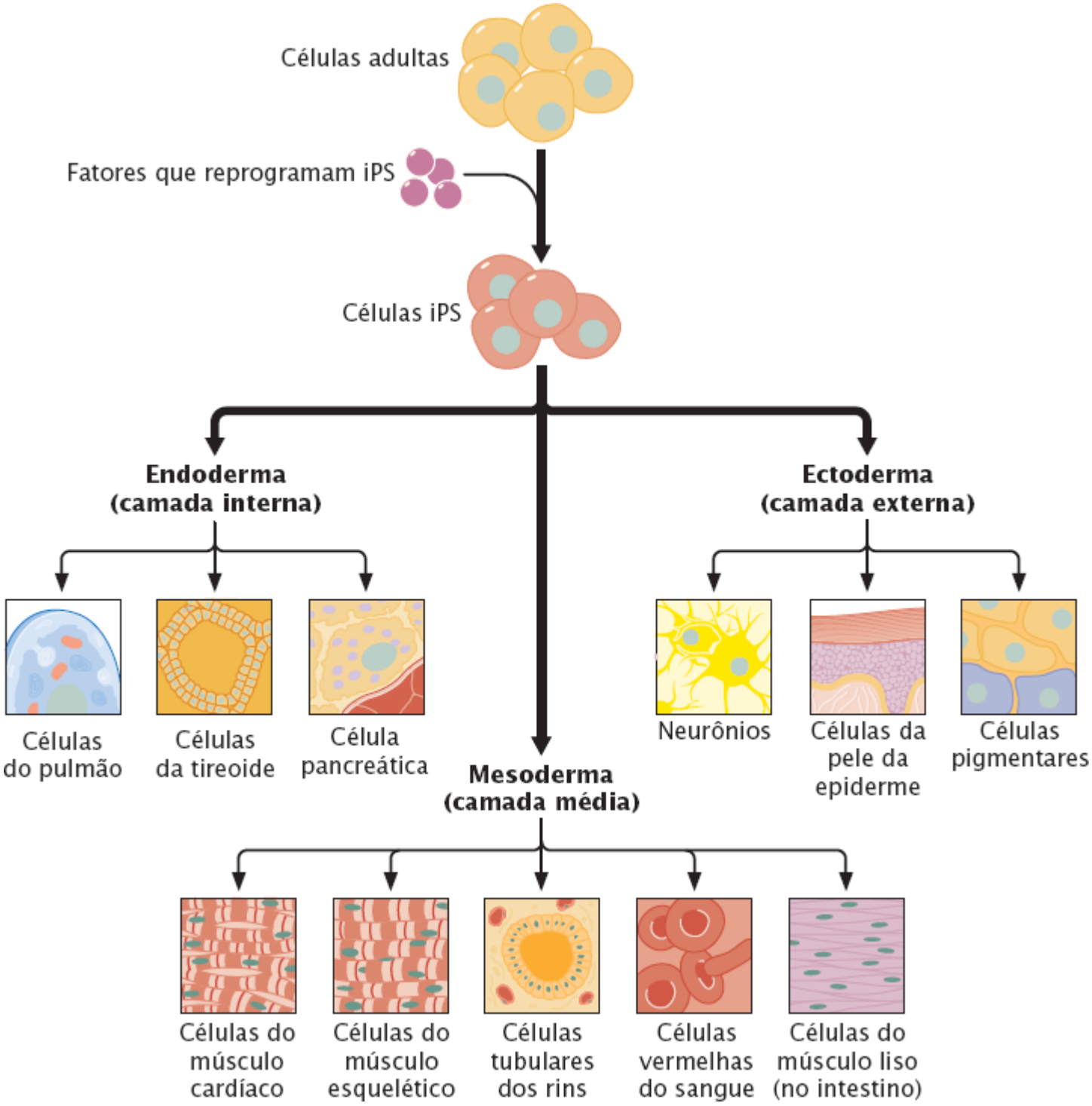


Figura 21.9 As células adultas diferenciadas podem ser reprogramadas para formar células-tronco pluripotentes induzidas (iPSCs), que conseguem se diferenciar em vários tipos de células.

Imprinting genômico

Os organismos diploides em geral têm dois alelos em cada *locus* autossômico, um alelo herdado da mãe e um alelo herdado do pai. Para a maioria dos genes, ambos os alelos são expressos, e o efeito de um alelo específico no fenótipo é independente de qual genitor transmitiu o alelo para os descendentes. Entretanto, para alguns genes, o sexo do genitor que contribuiu com o alelo influencia como este é expresso; alelos herdados da mãe e do pai não são equivalentes (**Figura 21.10**). Esse fenômeno, no qual o sexo do genitor que transmite o alelo determina sua expressão, é chamado de *imprinting* genômico (ver *Capítulo 5*). Para alguns genes imprintados (*imprinted*), o alelo herdado do genitor é expresso e o alelo herdado da genitora é silenciado; para outros genes, o alelo herdado da genitora é expresso e o alelo herdado do genitor é silenciado. Como foi discutido no Capítulo 5, acredita-se que o *imprinting* genômico seja responsável por diferentes graus de metilação dos genes herdados dos genitores.

A pesquisa mais antiga sugeria que o número de genes imprintados era limitado, mas estudos mais recentes sugerem que o número é muito maior. Um estudo conduzido por Christopher Gregg e seus colaboradores na Harvard University descobriu que mais de 1.300 genes no cérebro do camundongo exibiam indícios de *imprinting* genômico. Muitos desses genes imprintados não foram completamente silenciados: existe uma expressão tendenciosa, com um sexo transmitindo um alelo que era muito mais

expresso do que o alelo transmitido pelo outro sexo. Gregg e seus colaboradores também descobriram que o *imprinting* variava muito; alguns genes eram imprintados apenas em alguns tecidos ou em alguns momentos do desenvolvimento.

O *imprinting* genômico tem vários paralelos interessantes com a inativação do X. A maioria dos genes imprintados está localizada nos *clusters* de 3 a 12 genes que surgem em uma região bem definida de um cromossomo específico. Cada *cluster* tem genes codificadores de proteínas, assim como genes que produzem RNA não codificado. Em cada um dos exemplos mais bem estudados, existe uma região de controle de *imprinting* que determina o *imprinting*; a deleção dessa região destroi a capacidade de "imprintar". Além disso, a região de controle de *imprinting* exibe diferentes modificações de cromatina entre os alelos herdados do genitor e da genitora. Cada *cluster* de *imprinting* tem genes para um ou mais lncRNAs importantes no *imprinting* e eles próprios estão imprintedados. Por exemplo, o gene para o fator de crescimento 2 semelhante à insulina (*Igf2*) nos humanos exibe *imprinting* genômico, o alelo *Igf2* transmitido pelo genitor é expresso enquanto o alelo *Igf2* transmitido pela genitora é silenciado (ver Figura 5.18, no *Capítulo 5*). São necessários vários lncRNAs produzidos por outros genes na região de controle de *imprinting* para silenciar o gene de *Igf2* nas fêmeas, embora não esteja claro como eles produzem a repressão da transcrição.

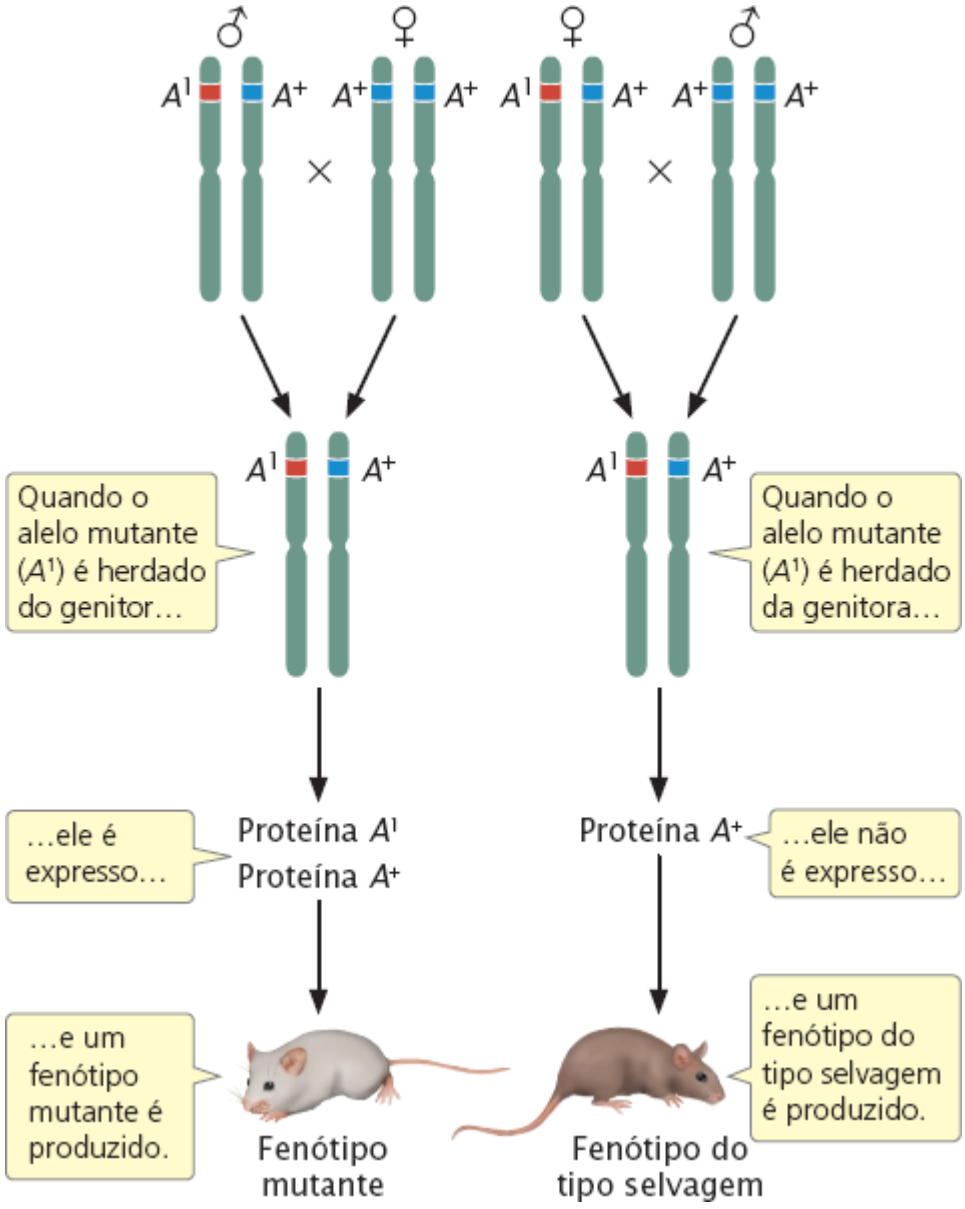


Figura 21.10 No *imprinting* genômico, a expressão de um alelo depende de ele ser herdado do genitor ou da genitora. Nesse caso, o alelo A^1 é expresso apenas quando herdado do genitor, mas, em outros casos, um alelo é expresso apenas quando herdado da genitora.

Muitos dos *clusters* bem estudados de genes imprintados estão associados a distúrbios resultantes de *imprinting* defeituoso. A síndrome de Beckwith-Wiedemann é um destes distúrbios. As crianças com essa síndrome exibem crescimento excessivo durante o desenvolvimento fetal e a primeira infância. Eles também têm tumores malignos embrionários incomuns. A síndrome de Beckwith-Wiedemann está associada ao *imprinting* de um *cluster* de genes no cromossomo 11, incluindo o gene *Igf2*. Os indivíduos com essa síndrome têm pequenas deleções no cromossomo 11 que interferem no processo normal de *imprinting*. Por

exemplo, *Igf2* é normalmente expresso apenas quando herdado do pai, mas em algumas crianças com a síndrome de Beckwith-Wiedemann, as deleções dentro do centro de *imprinting* podem levar à expressão de alelos de ambos os genitores. O resultado é a produção em grande quantidade de *Igf2*, levando ao crescimento excessivo e câncer. A síndrome de Prader-Willi e a síndrome de Angelman são distúrbios causados por defeitos no *imprinting* no cromossomo 15 (ver *Capítulo 15*).

Imprinting e conflito genético. Muitos genes que são genomicamente imprintados afetam o crescimento fetal e embrionário inicial. Uma possível explicação para o *imprinting* genômico é a **hipótese de conflito genético**, que sugere que existem pressões evolutivas diferentes e conflitantes agindo sobre os alelos maternos e paternos para genes (como o *Igf2*) que afetam o crescimento fetal. A partir do ponto de vista evolutivo, os alelos paternos que maximizam o tamanho dos descendentes são favorecidos, porque o peso ao nascer está fortemente associado à mortalidade infantil e à saúde do adulto. Assim, é vantajoso que o genitor transmita os alelos que promovem o crescimento fetal máximo dos seus descendentes. Em contrapartida, os alelos maternos que causam crescimento fetal mais limitado são favorecidos: entregar uma quantidade muito excessiva de nutrientes para qualquer feto pode limitar a capacidade da genitora em se reproduzir no futuro e dar à luz fetos macrossômicos também é difícil e arriscado. A hipótese de conflito genético prevê que o *imprinting* genômico se desenvolverá: as cópias paternas dos genes que afetam o crescimento fetal serão expressas ao máximo, enquanto as cópias maternas dos mesmos genes serão expressas de forma menos ativa ou até silenciadas. De fato, o *Igf2* segue este padrão: o alelo paterno está ativo e promove o crescimento, o alelo materno é silenciado e não contribui com o crescimento. Achados recentes demonstram que a cópia paterna de *Igf2* promove o crescimento fetal ao direcionar mais nutrientes maternos para o feto através da placenta.

Conceitos

O *imprinting* genômico é causado por diferenças epigenéticas nos alelos herdados dos genitores. A hipótese de conflito genético sugere que o *imprinting* se desenvolve por causa das pressões evolutivas conflitantes que agem sobre os alelos maternos e paternos.

Checagem dos conceitos 6

O que é verdade sobre o *imprinting* genômico?

a. O sexo do genitor que transmite um alelo afeta a expressão do alelo nos descendentes.
b. O sexo dos descendentes afeta a expressão de um alelo herdado de um dos genitores.
c. O sexo do genitor afeta como um alelo é transmitido para os descendentes.
d. O sexo dos descendentes afeta qual alelo é herdado do genitor.

21.4 Epigenoma

Em 2003, pesquisadores declararam que praticamente o genoma humano inteiro tinha sido sequenciado. Esse feito monumental foi extremamente valioso no esclarecimento de como as informações genéticas são codificadas no genoma. Todavia, a sequência de bases do DNA é apenas um registro parcial das informações hereditárias. Como já discutimos, a estrutura da cromatina tem informações epigenéticas adicionais: informações que são herdadas e afetam como a sequência de bases é expressa. O padrão geral das modificações da cromatina de cada organismo é chamado de **epigenoma**. Nos últimos anos, foram disponibilizadas técnicas para detectar e descrever modificações epigenéticas pelo genoma.

Detecção da metilação do DNA. Foram desenvolvidas várias técnicas para examinar os níveis de metilação de DNA. Algumas dessas dependem de endonucleases de restrição, enzimas que fazem cortes de fita dupla no DNA em sequências de bases específicas (ver *Capítulo 19*). Algumas enzimas de restrição são sensíveis à metilação e não cortarão uma sequência que tenha 5-metilcitosina, enquanto outras enzimas de restrição são insensíveis à metilação. Pode-se determinar os padrões gerais de metilação ao cortar o DNA com enzimas que são sensíveis à metilação e com enzimas que não são e então analisar os fragmentos resultantes.

Uma técnica muito usada e mais precisa para analisar a metilação do DNA é o sequenciamento pelo bissulfito (**Figura 21.11**). Nessa técnica, o DNA genômico é tratado primeiro com bissulfito de sódio, que converte quimicamente a citosina não metilada em uracila. As uracilas são então detectadas como timina durante o sequenciamento. Entretanto, a 5-metilcitosina não é quimicamente alterada pelo tratamento com bissulfito e é detectada como citosina durante o sequenciamento (ver *Capítulo 19* para uma discussão sobre o sequenciamento do DNA). Ao sequenciar o DNA genômico com e sem o tratamento com bissulfito, os pesquisadores podem determinar as localizações de todas as cópias de 5-metilcitosina no DNA.

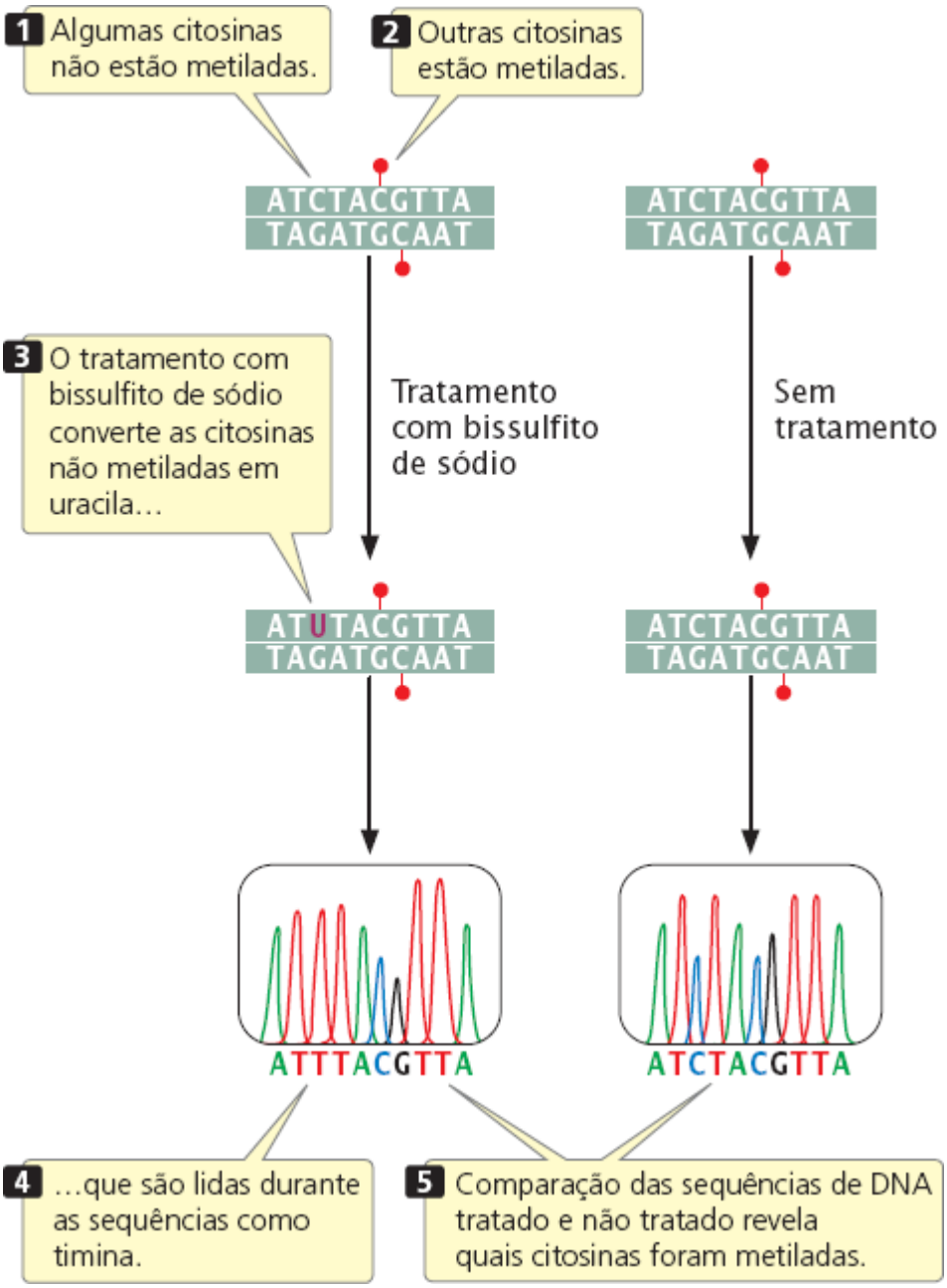


Figura 21.11 O sequenciamento por bissulfito pode ser usado para determinar as localizações das 5-metilcitosinas.

Detecção das modificações de histona. As modificações de histona podem ser detectadas por fragmentação da cromatina e aplicação de um anticorpo específico para uma modificação específica da histona, um processo chamado imunoprecipitação da cromatina (abreviado como ChIP, ver *Capítulo 17*). O anticorpo provoca a precipitação da cromatina com a modificação da histona e consequente separação dos fragmentos de cromatina sem modificação. A proteína é então removida por digestão com uma enzima que a degrada, mas não o DNA, e o fragmento de DNA com o qual a histona estava associada é sequenciado. Esse processo fornece informações sobre onde ocorrem as modificações de histona no genoma.

Marcadores epigenéticos de genoma inteiro. Usando essas técnicas, os geneticistas compararam os epigenomas de diferentes tipos de células. Por exemplo, os pesquisadores determinaram a distribuição de 5-metilcitosina pelo genoma inteiro em dois tipos de células: (1) uma célula-tronco humana não diferenciada e (2) um fibroblasto. Os pesquisadores descobriram diferenças comuns nos padrões de metilação dessas células. O achado que os padrões de metilação de DNA varia entre os tipos de células apoia a ideia de que a metilação da citosina fornece um meio pelo qual os tipos celulares mantêm suas diferenças durante o desenvolvimento. Outros pesquisadores compararam os epigenomas de células de câncer e células normais e observaram marcadores epigenéticos distintos associadas ao câncer.

Da mesma forma, os pesquisadores mapearam as localizações genômicas de modificações de histona em diferentes tipos celulares. Estes estudos detectaram modificações de histona específicas associadas a promotores e acentuadores de genes ativos. Em um estudo, os pesquisadores mapearam nove marcadores epigenéticos diferentes em nove tipos diferentes de células

humanas. Eles conseguiram determinar como as marcas de cromatina variavam entre os tipos celulares e compararam os marcadores epigenéticos associados aos genes ativos e reprimidos. Como os marcadores epigenéticos específicos estão associados a elementos reguladores como promotores e acentuadores, os pesquisadores usaram esses marcadores para mapear as localizações desses elementos reguladores por todo genoma.

Conceitos

O epigenoma é o conjunto completo de modificações da cromatina de um organismo.

Resumo dos conceitos

- O termo epigenética, criado por Conrad Waddington, atualmente refere-se aos efeitos pelos quais os fenótipos são transferidos para outras células ou gerações futuras sem diferenças nas sequências de bases do DNA
- Muitos fenótipos epigenéticos resultam de alterações na estrutura da cromatina. Os efeitos epigenéticos ocorrem por meio da metilação do DNA, modificação da histona e moléculas de RNA
- As mudanças epigenéticas são estáveis, mas são potencialmente afetadas por fatores ambientais
- Três mecanismos moleculares sustentam muitos fenótipos epigenéticos: (1) mudanças nos padrões de metilação do DNA; (2) modificações químicas das proteínas histona e (3) moléculas de RNA que afetam a estrutura da cromatina e a expressão gênica
- Alguns efeitos epigenéticos surgem a partir da metilação do DNA, na qual as bases citosina são metiladas para formar a 5-metilcitosina. A metilação ocorre em dinucleotídios CpG. A presença de 5-metilcitosina está associada à repressão da transcrição
- Algumas regiões do DNA com muitos dinucleotídios CpG são chamadas de ilhas de CpG. A metilação das ilhas de CPG próximas de um gene leva à repressão da transcrição
- A metilação do DNA inibe a transcrição ao inibir a ligação dos fatores de transcrição e outras proteínas necessárias para que ocorra a transcrição. A metilação do DNA também atrai proteínas que reprimem a transcrição e enzimas desacetilases de histona que alteram a estrutura da cromatina
- A metilação do DNA é mantida durante a divisão celular pelas enzimas metiltransferases que reconhecem a metilação dos dinucleotídios CpG em uma fita de DNA e adicionam grupos metila a bases citosina não metiladas na outra fita
- As modificações das proteínas histona alteram a estrutura da cromatina. As modificações da histona podem ser transmitidas pela divisão celular
- As moléculas de RNA modificam a cromatina por meio de vários processos
- A paramutação é uma alteração hereditária de um alelo por outro alelo, sem qualquer mudança na sequência de DNA. A paramutação no milho e nos camundongos é mediada por pequenas moléculas de RNA
- As experiências no início da vida podem produzir mudanças epigenéticas que têm efeitos duradouros sobre o comportamento
- As substâncias químicas do ambiente podem produzir efeitos epigenéticos que são transmitidos para as gerações posteriores
- As modificações epigenéticas têm efeitos sobre o metabolismo que se estendem pelas gerações
- As diferenças fenotípicas entre os gêmeos monozigotos geneticamente idênticos podem surgir a partir dos efeitos epigenéticos
- A inativação de X ocorre quando um cromossomo X nas células das fêmeas é silenciado de forma permanente. As mudanças epigenéticas produzem a inativação do X e requerem a ação de vários genes que codificam longos RNAs não codificados
- O *imprinting* genômico ocorre quando a expressão de um gene depende de qual genitor transmite o gene. Ele é causado por mudanças epigenéticas na estrutura de cromatina que são transmitidas para os descendentes. A hipótese de conflito genético sugere que o *imprinting* se desenvolve por causa das pressões evolutivas conflitantes que agem sobre os alelos maternos e paternos
- O epigenoma é o conjunto completo de modificações da cromatina de um organismo.

Termos importantes

Células-tronco
Células-tronco pluripotentes induzidas
Centro de inativação do X
Epialelos
Epigenoma
Hipótese de conflito genético
Marcadores epigenéticos

Paramutação
Pluripotência

Respostas da Checagem dos conceitos

1. b.
2. b.
3. d.
4. c.
5. Não haveria RNA *Xist* para recobrir o cromossomo X e a inativação não ocorreria. Ambos os cromossomos X permaneceriam ativos.
6. a.

Problema desenvolvido

Problema

O alelo *b1* codifica um fator de transcrição que estimula a produção de antocianina, um pigmento púrpura nas plantas. Qual seria o efeito de eliminar sete repetições em *tandem* localizadas 100.000 pb *upstream* do *locus b1* no milho?

Estratégias para a solução

De qual informação você precisa para solucionar o problema?
Os efeitos da deleção das repetições.

Quais informações são fornecidas para solucionar o problema?
- O *locus b1* codifica um fator de transcrição que estimula a produção de antocianina nas plantas
- As sete repetições em *tandem* estão *upstream* do *locus* b1.

Para a solução deste problema, revise:
Paramutação no milho, na Seção 21.3.

Etapas para a solução

As informações fornecidas no texto indicam que a planta *B-I B-I* normalmente tem alta expressão do *locus b1*, produz antocianina e tem cor púrpura. As repetições em *tandem* são necessárias para a elevada transcrição do *locus b1*, que codifica um fator de transcrição, responsável pela produção de antocianina, um pigmento púrpura. As repetições em *tandem* atuam como um acentuador, estimulando a transcrição do *locus b1* no genótipo *B-I B-I*. Sem a ação semelhante de acentuador das repetições, *b1* será transcrito em níveis mínimos e pouca antocianina será produzida. Esse fenômeno resultará em plantas *B-I B-I* com poucos pigmentos, o mesmo fenótipo observado nas plantas com genótipo *B′B′*.

Lembrete: Os acentuadores estimulam a transcrição em genes distantes deles (ver *Capítulo 17*).

As repetições em *tandem* também codificam 25 siRNAs com 25 nucleotídios de comprimento, necessários para a paramutação (conversão do alelo *B-I* em alelos *B′**). Assim sendo, a deleção das repetições em *tandem* eliminaria a capacidade dos alelos *B′* em fazer paramutação.

Questões de compreensão

Introdução

1. O que é a hipótese do fenótipo poupador? Como ela ajuda a explicar os efeitos a longo prazo da dieta observados nos residentes de Överkalix?

Seção 21.1

*2. Quais são as características importantes de um traço epigenético?

Seção 21.2

3. Quais são os três mecanismos moleculares que alteram a estrutura da cromatina e responsáveis por muitos fenótipos epigenéticos?
4. Qual é a principal forma de metilação de DNA observada nos eucariotos? Em qual tipo de sequência de DNA é comum encontrar a metilação do DNA?
5. Como a metilação do DNA reprime a transcrição?
6. Explique resumidamente como os padrões de metilação do DNA são transmitidos pela divisão celular.
7. Quais são os tipos de modificações de histona responsáveis pelos fenótipos epigenéticos?

Seção 21.3

8. O que é paramutação? Quais são as características fundamentais deste fenômeno?
9. Descreva resumidamente a paramutação no *locus Kit* de camundongos. Que evidências sugerem que pequenas moléculas de RNA são importantes para esse fenômeno?
10. Que evidências sugerem que a cognição nos camundongos é influenciada por mudanças epigenéticas?
11. Explique como a vinclozolina atua como desregulador endócrino.
12. Dê um exemplo de efeito epigenético transgeração da dieta sobre o metabolismo.
13. Que evidências sugerem que as diferenças nos gêmeos monozigóticos podem ser causadas por efeitos epigenéticos?
14. Como a inativação do X é um fenótipo epigenético?
15. Descreva resumidamente os processos moleculares que fazem um cromossomo X em cada célula de fêmea ser ativado e o outro cromossomo X ser inativado.
16. O que são células-tronco pluripotentes induzidas? Como elas são derivadas das células somáticas adultas?
17. Defina *imprinting* genômico.
18. O que é hipótese de conflito genômico para a origem do *imprinting* genômico.

Seção 21.4

19. O que é epigenoma?

Questões e problemas aplicados

Introdução

20. A introdução deste capítulo descreve os efeitos duradouros da dieta sobre os residentes de Överkalix, Suécia.
    a. Que evidências sugerem que esses efeitos são causados por efeitos epigenéticos?
    b. Que evidência adicional ajudaria a demonstrar que essas mudanças são causadas por mudanças epigenéticas?

Seção 21.1

21. Qual é a diferença dos traços epigenéticos dos traços genéticos tradicionais, como as diferenças na cor e no formato das ervilhas que Mendel estudou?
22. A epigenética é descrita como “herança, mas não como a conhecemos”. Você acha essa definição adequada? Por que sim ou por que não?

Seção 21.2

23. O que seria necessário para provar que um fenótipo é causado por uma mudança epigenética?
24. Qual abelha da Figura 21.3 (a operária ou a rainha) terá mais cópias de 5-metilcitosina no seu DNA? Explique sua resposta.

*25. Qual seria o efeito de eliminar o gene *Dnmt3* nas abelhas?

*26. Parte da metilação do DNA nos eucariotos ocorre nos dinucleotídios CpG, mas alguns nucleotídios individuais de citosina também são metilados para formar a 5-metilcitosina. Considerando o que você sabe sobre o processo pelo qual a metilação do DNA nos dinucleotídios CpG é mantida durante a divisão celular, você acha que a metilação dos nucleotídios C também seria mantida pelo mesmo processo? Explique sua justificativa.

Seção 21.3

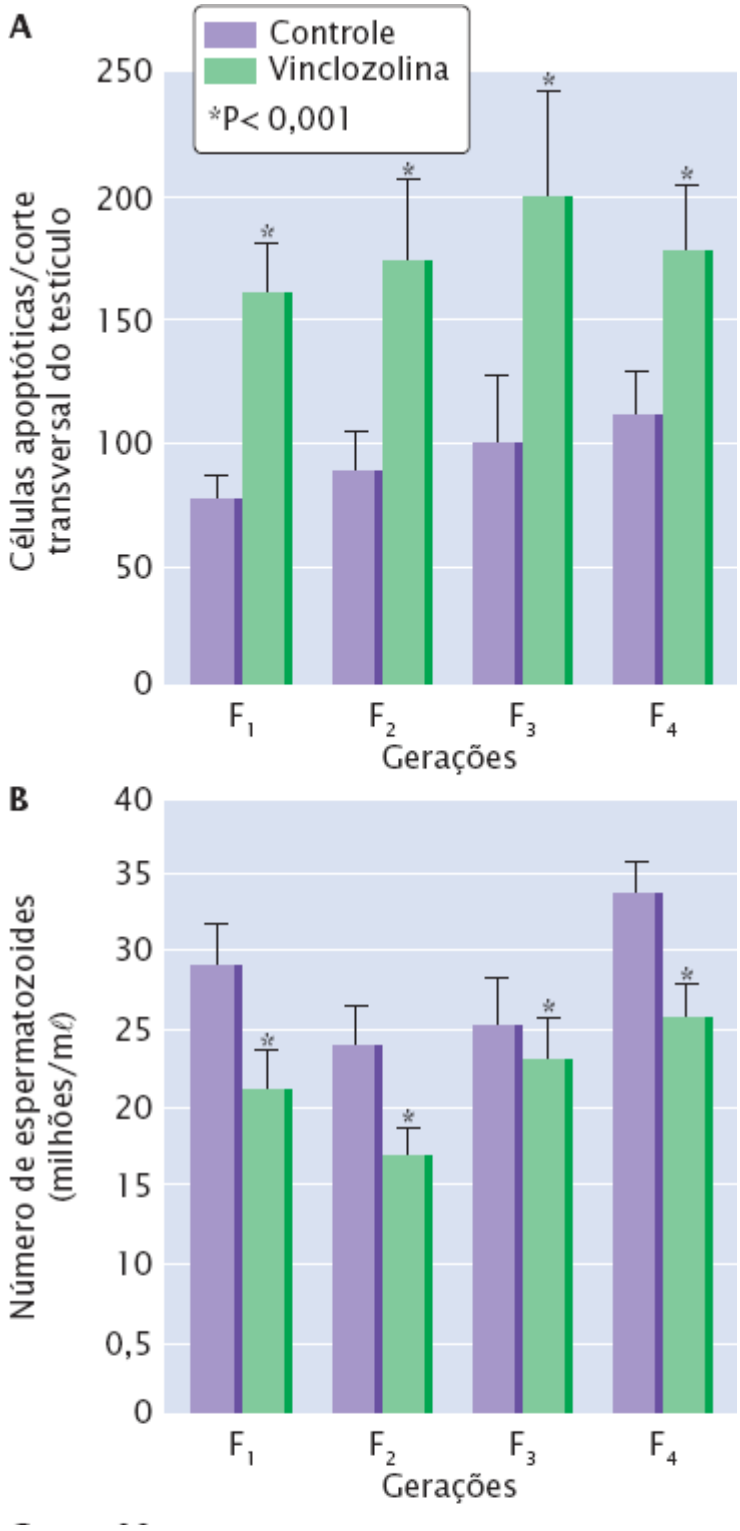
27. Um cruzamento entre o indivíduo F_1 na Figura 21.4 e uma planta com genótipo *B-I B-I* produzirá descendentes com qual fenótipo?

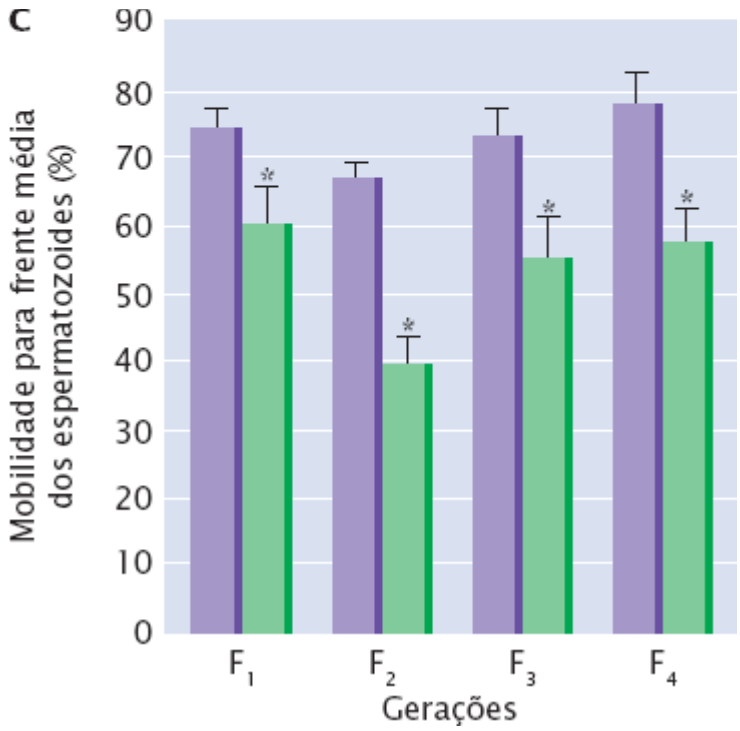
*28. Uma cientista faz um experimento no qual remove os descendentes dos ratos da mãe ao nascer e pede a seus estudantes de genética que alimentem e criem os filhotes. Considerando que os estudantes não lambem nem cuidam desses filhotes como as mães normalmente o fazem, quais seriam os efeitos comportamentais e epigenéticos a longo prazo esperados

nesses ratos quando eles crescerem?

ANÁLISE DE DADOS

*29. Ratas prenhes foram expostas a uma dose diária de 100 ou 200 mg/kg de vinclozolina, um fungicida muito usado na indústria de vinho (M. D. Anway *et al*. 2005. *Science* 308:1466-1469). Os descendentes F_1 das fêmeas expostas foram intercruzados, produzindo os ratos F_2, F_3 e F_4. Observe que nenhum dos ratos F_2, F_3 e F_4 foi exposto à vinclozolina. Os testículos dos ratos machos F_1-F_4 foram examinados e comparados com os dos ratos de fêmeas do controle que não foram expostas à vinclozolina (ver os gráficos anteriores). Esses efeitos foram observados em mais de 90% dos descendentes machos de F_1-F_4. Além disso, 8% dos machos F_1-F_4 das fêmeas expostas à vinclozolina desenvolveram infertilidade completa, comparados com 0% dos machos F_1-F_4 das fêmeas do controle. A análise molecular dos testículos demonstrou que os padrões de metilação do DNA eram diferentes entre os descendentes das fêmeas expostas à vinclozolina dos descendentes das fêmeas do controle. Forneça uma explicação para os efeitos transgeração da vinclozolina sobre a fertilidade dos machos.





(Gráficos de M. D. Anway *et al.* 2005. *Science* 308:1466-1469.)

30. Com base nas informações de estudos de efeitos duradouros da dieta sobre o metabolismo nos camundongos, quais poderiam ser os efeitos epigenéticos nos filhos e netos das pessoas de Överkalix que foram expostas a períodos de fome extrema na infância? Inclua na sua resposta os tipos de mudanças epigenéticas à cromatina que você pode esperar ver e os efeitos fenotípicos no metabolismo dos lipídios e do colesterol.

*31. Qual seria o efeito sobre a inativação do X ao adicionar siRNAs que eliminariam os produtos de cada um dos seguintes genes?
   a. *Xist*.
   b. *Jpx*.

Seção 21.4

32. Um fragmento de DNA com a seguinte sequência de bases tem algumas bases citosina que são metiladas (indicadas por C*) e outras que são não metiladas. Para determinar a localização de citosinas metiladas e não metiladas, os pesquisadores sequenciaram esse fragmento com e sem o tratamento com bissulfito de sódio. Indique a sequência de bases que será lida com e sem o tratamento com bissulfito.

— ATCGC*GTTAC *ATTGC *ATCA —

33. Uma geneticista está interessada em determinar as localizações das citosinas metiladas em um fragmento de DNA. Ela trata algumas cópias do fragmento com bissulfito de sódio e deixa algumas cópias sem tratamento. Depois sequencia as cópias tratadas e não tratadas do fragmento e obtém os seguintes resultados. Indique a sequência original do fragmento de DNA e indique a presença das citosinas metiladas.

Sequência
sem tratamento: — AATTGCCCGATCGATTAAGCCA —
Sequência
com tratamento: — AATTGTTTGATCGATTAAGCTA —

Questões desafiadoras

Seção 21.3

34. É provável que a hipótese de conflito genômico explique o *imprinting* genômico para os genes envolvidos na memória dos adultos? Por que sim ou por que não?

ANÁLISE DE DADOS

35. Nos últimos anos, foram desenvolvidas técnicas para clonar mamíferos por um processo chamado de transferência nuclear, na qual o núcleo de uma célula somática é transferido para um óvulo do qual foi removido o material nuclear (ver *Capítulo 22*). A pesquisa demonstrou que quando um núcleo de uma célula somática diferenciada é transferido para um óvulo, apenas uma pequena porcentagem dos embriões resultantes completa o desenvolvimento e muito deles morrem pouco depois de nascer. Em contrapartida, quando um núcleo de uma célula-tronco embrionária não diferenciada é transferido para um óvulo, a porcentagem de embriões que completam o desenvolvimento é muito maior (W. M. Rideout, K. Eggan, and R. Jaenisch. 2001. *Science* 293:1095-1098). Proponha uma possível explicação para o desenvolvimento bem-sucedido de embriões clonados ser maior quando o núcleo transferido provém de uma célula-tronco embrionária não diferenciada.

Seção 21.4

36. Foi proposto o uso de células-tronco embrionárias para substituir células destruídas por doença ou lesão. Por causa de questões éticas sobre a criação e a destruição de embriões para produzir células-tronco embrionárias, os pesquisadores tentaram criar células pluripotentes induzidas (iPSCs). Neste capítulo, discutimos estudos que mostram que iPSCs conservam alguns marcadores epigenéticos das células adultas diferenciadas de onde derivaram. Quais implicações essa pesquisa poderia ter para as tentativas de usar iPSCs para cultivo de células e tecidos?

22

Genética Desenvolvimental e Imunogenética

Origem dos esgana-gatas sem espinhas dorsais e pélvicas

Os esgana-gatas (*Gasterosteus aculeatus*) são pequenos peixes curiosos. Apesar do seu tamanho pequeno – apenas 5 cm de comprimento –, eles têm pesada carapaça corporal, placas de proteção nas costas, nas laterais e no ventre e três espinhas na sua superfície dorsal. Cada peixe tem espinhas pélvicas impressionantes, que estão ancoradas na cintura pélvica e se projetam para suas laterais. As espinhas dorsais e pélvicas dificultam engolir esse peixe, permitindo que eles sobrevivam em um ambiente onde um peixe de 5 cm desprotegido seria uma refeição fácil para numerosos predadores.

A maioria dos esgana-gatas é marinha, mas algumas populações isoladas podem ser encontradas em lagos de água doce. Na América do Norte, essas populações de água doce se originaram há 10.000 a 20.000 anos atrás, no final da última era do gelo, quando os esgana-gatas marinhos invadiram os lagos. Muitas populações lacustres dos esgana-gatas perderam sua carapaça e suas espinhas, provavelmente porque não existem os predadores do peixe que poderiam comê-los, existe pouco cálcio para desenvolver as placas e espinhas e os predadores invertebrados encontrados no lago pegam o peixe ao se agarrarem a suas espinhas. Os biólogos estão interessados em como os esgana-gatas marinhos fizeram a transição evolutiva do ambiente marinho para a água doce: como um peixe com pesada carapaça ficou sem espinhas? A pesquisa feita pelos geneticistas do desenvolvimento começou a responder a essa pergunta.

Em 1998, o geneticista David Kingsley da Stanford University começou a trabalhar com Dolph Schluter, um biólogo evolucionista da University of British Columbia. Seu objetivo era compreender como os esgana-gatas lacustres perderam suas espinhas pélvicas durante a transição evolutiva dos ambientes de água salgada para água doce. Os cientistas cruzaram um esgana-gata marinho fêmea que tinha espinhas com um macho do Paxton Lake, British Columbia, que não tinha espinhas pélvicas. Todos os peixes desse cruzamento tinham espinhas pélvicas. Eles então cruzaram dois peixes de F_1, produzindo um total de 375 descendentes F_2. Esses descendentes F_2 apresentavam uma grande faixa de variação nas suas espinhas pélvicas, alguns tinham espinhas totalmente desenvolvidas, alguns não tinham espinhas e outros tinham graus variáveis de redução das espinhas.

Kingsley e colaboradores examinaram a associação de espinhas pélvicas nos descendentes F_2 e a herança de marcadores genéticos pelo genoma. Eles descobriram que a maioria das variações nas espinhas pélvicas estava associada com marcadores genéticos de uma região específica no cromossomo 7. E interessante, essa mesma região tem *Pitx1*, um gene expresso nos membros posteriores dos vertebrados em desenvolvimento. Camundongos com uma mutação no seu gene *Pitx1* têm membros posteriores reduzidos, assim como outras anomalias de desenvolvimento. Essa observação sugeriu que as mutações no gene *Pitx1* podem ser responsáveis pela ausência de espinhas pélvicas nas populações de água doce de esgana-gatas. Mas, quando os pesquisadores examinaram as sequências de DNA do gene *Pitx1* nos esgana-gatas de água salgada e água doce, não encontraram diferenças que alterariam a sequência de aminoácidos da proteína codificada pelo gene. Eles descobriram que a expressão do gene *Pitx1* era diferente entre esses tipos de peixes. Os esgana-gatas com barbatanas pélvicas expressavam *Pitx1* na sua região pélvica, os peixes sem as barbatanas pélvicas expressavam esse gene em outros tecidos, mas o gene estava totalmente inativo na região pélvica. Essa descoberta sugeria que, embora o gene *Pitx1* por si só não fosse mutante nos esgana-gatas, as mutações nos elementos reguladores da sua expressão poderiam ser a fonte da variação na existência ou não de espinhas pélvicas.

Em 2010, Kingsley e sua equipe localizaram a mutação que causa a ausência de espinhas nas populações de água doce dos esgana-gatas. Eles encontraram um acentuador 500 pb *upstream* do gene *Pitx1* que controla a expressão específica de tecido do *Pitx1*. Os acentuadores são sequências de DNA que promovem a transcrição de genes distantes, em geral específicos por tecidos (ver *Capítulo 17*). Os pesquisadores determinaram que o peixe sem espinha do Paxton Lake tinha uma deleção que removeu esse acentuador, assim evitando a expressão de *Pitx1* e por fim levando à perda das espinhas pélvicas. Os peixes sem espinha de outros lagos do Canadá, Alasca e até Islândia também tinham deleções do acentuador, mas os peixes de diferentes lagos tinham

diferentes deleções. Essa observação sugere que durante o curso da evolução as espinhas pélvicas foram perdidas múltiplas vezes pela seleção natural agindo em diferentes mutações que têm o mesmo efeito fenotípico.

A história de como o esgana-gatas perdeu suas espinhas ilustra que podem ocorrer grandes alterações anatômicas por meio de pequenas mudanças genéticas nas sequências reguladoras importantes que afetam o desenvolvimento, um tema discutido neste capítulo do controle genético do desenvolvimento. O capítulo começa com uma análise de como ocorre a diferenciação celular, não pela perda de genes, mas pela alteração da expressão gênica. Então vamos discutir o controle genético do desenvolvimento inicial dos embriões da *Drosophila*, um dos sistemas de desenvolvimento mais bem compreendidos. A seguir vamos analisar o controle genético da estrutura floral nas plantas, outro sistema de modelo bem estudado, seguida por uma observação mais detalhada na morte celular programada e o uso do desenvolvimento para compreender a evolução. Vamos estudar o desenvolvimento da imunidade e seu controle genético ao final do capítulo.

22.1 O desenvolvimento se dá por diferenciação celular

Cada organismo multicelular começa como um ovo unicelular, fertilizado. Esse zigoto unicelular sofre repetidas divisões celulares, produzindo milhões ou trilhões de células que constituem um organismo adulto completo. Inicialmente, cada célula no embrião é dotada de **totipotência**: tem o potencial para se desenvolver em qualquer tipo de célula. Muitas células nas plantas e fungos permanecem totipotentes, mas as células animais em geral ficam comprometidas em se desenvolver em tipos específicos de células logo após algumas divisões embrionárias. Esse comprometimento vem logo após a célula começar a exibir qualquer característica de um tipo celular específico; quando a célula se torna comprometida, normalmente ela não deixa de sê-lo nem se torna um tipo celular diferente. Uma célula se torna comprometida por um processo chamado **determinação**.

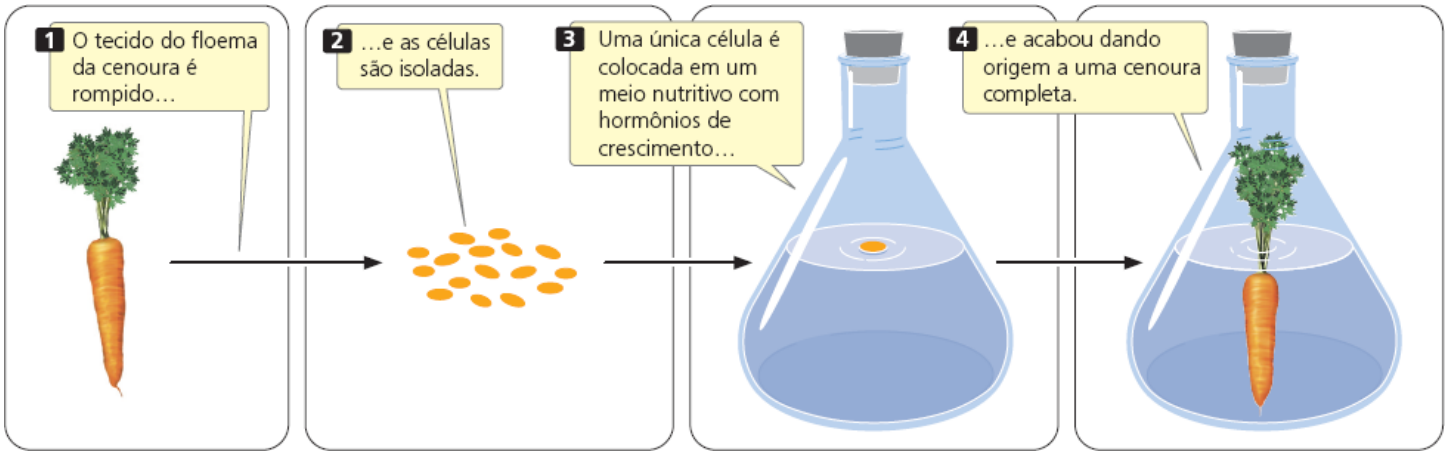


Figura 22.1 Muitos tipos de plantas podem ser clonados a partir de células isoladas. Este tipo de experimento demonstra que nenhum material genético original é perdido durante o desenvolvimento da planta.

Por muitos anos, o trabalho dos biólogos do desenvolvimento estava limitado em descrever as mudanças que ocorriam durante o desenvolvimento, porque não existiam técnicas para sondar os processos intracelulares subjacentes. Mas, nos últimos anos, poderosas técnicas genéticas e moleculares tiveram uma grande influência no estudo do desenvolvimento; por exemplo, o sequenciamento de DNA forneceu muitas informações sobre a natureza e a organização das sequências de DNA que controlam os processos do desenvolvimento. Em alguns sistemas modelo como *Drosophila* e *Arabidopsis*, os mecanismos moleculares envolvidos na mudança do desenvolvimento estão começando a ser compreendidos.

Se todas as células em um organismo multicelular provêm da mesma célula original, como surgem os diferentes tipos de células? Antes dos anos 1950, existiam duas hipóteses. Uma possibilidade era que os genes de desenvolvimento poderiam ser seletivamente perdidos ou alterados, fazendo com que os diferentes tipos celulares tenham diferentes genomas. Por outro lado, cada célula poderia ter as mesmas informações genéticas, mas diferentes genes poderiam ser expressos em cada tipo de célula. Os resultados dos primeiros experimentos de clonagem ajudaram a resolver essa questão.

Experimentos de clonagem em plantas

Nos anos 1950, Frederick Steward desenvolveu métodos para clonar plantas. Ele rompeu o floema (tecido condutor) da raiz de uma cenoura ao separar e isolar células individuais e depois colocou as células em um meio estéril com nutrientes e outras substâncias necessárias para crescer. Steward conseguiu fazer com que as células crescessem e se dividissem e acabou obtendo cenouras comestíveis a partir de células isoladas (**Figura 22.1**). Como todas as partes da planta foram regeneradas de uma célula especializada do floema, ele concluiu que cada célula do floema tinha o potencial genético para uma planta inteira;

nenhum material genético original foi perdido durante a determinação.

Experimentos de clonagem em animais

Os resultados de outros estudos demonstraram que a maioria das células dos animais também conserva um conjunto completo de informações genéticas durante o desenvolvimento. Em 1952, Robert Briggs e Thomas King removeram os núcleos de ovócitos não fertilizados do sapo *Rana pipiens*. Depois isolaram os núcleos de blástulas de sapo (um estágio embrionário inicial), injetaram esses núcleos individualmente nos ovócitos e picaram os ovos com uma agulha para estimulá-los a se dividirem. Embora a maioria fosse danificada no processo, alguns ovos se desenvolveram em girinos completos que se tornaram sapos.

No final dos anos 1960, John Gurdon usou esses métodos para clonar com sucesso alguns sapos com núcleos isolados das células intestinais de girinos. Essa experiência sugeriu que as células intestinais diferenciadas carreavam as informações genéticas necessárias para codificar os traços encontrados em todas as outras células. Entretanto, as clonagens bem-sucedidas de Gurdon resultaram da existência de algumas células-tronco não diferenciadas no tecido intestinal, que foram usadas inadvertidamente como doadoras de núcleos.

Em 1997, pesquisadores do Roslin Institute of Scotland anunciaram que haviam conseguido clonar com sucesso uma ovelha ao usar o material genético de uma célula diferenciada de animal adulto. Para fazer esse experimento, usaram uma célula da teta de uma ovelha Finn Dorset com um ovo sem núcleo e estimularam eletricamente o ovo para iniciar o desenvolvimento. Após cultivar o embrião no laboratório por uma semana, eles o implantaram na mãe substituta, uma ovelha da raça Blackface escocesa. Dolly, o primeiro mamífero clonado a partir de uma célula adulta, nasceu no dia 5 de julho de 1996 (**Figura 22.2**). Desde então, vários outros animais, incluindo ovelhas, cabras, camundongos, coelhos, vacas, porcos, cavalos, mulas, cães e gatos foram clonados a partir de células adultas diferenciadas. É importante mencionar que, embora Dolly e outros mamíferos que tenham sido clonados tenham o mesmo material genético *nuclear* do seu genitor clonado, seus genes *citoplasmáticos* não são idênticos, como o cromossomo mitocondrial, porque o citoplasma é doado tanto pela célula doadora quanto pelo ovo sem núcleo.

Figura 22.2 Em 1996, pesquisadores do Roslin Institute of Scotland conseguiram clonar uma ovelha chamada Dolly. Eles usaram o material genético de uma célula diferenciada de animal adulto. (Paul Clements/AP.)

Os experimentos de clonagem demonstraram que o material genético não é perdido nem modificado de forma permanente durante o desenvolvimento: o desenvolvimento exige a expressão seletiva dos genes. Mas como as células regulam sua expressão gênica de forma coordenada para dar origem a um organismo complexo, multicelular? Agora a pesquisa começa a obter algumas respostas para essa pergunta crucial.

Conceitos

A capacidade de clonar plantas e animais a partir de células únicas especializadas demonstra que os genes não estão perdidos nem modificados de forma permanente durante o desenvolvimento.

Checagem dos conceitos 1

Os cientistas clonaram alguns animais ao injetar um núcleo de um embrião inicial em um ovo sem núcleo.

Como esse desfecho demonstra que o material genético não é perdido durante o desenvolvimento? Por que sim ou por que não?

22.2 A formação do padrão em *Drosophila* serve como modelo para o controle genético do desenvolvimento

A formação do padrão consiste nos processos do desenvolvimento que resultam no formato e na estrutura de organismos complexos, multicelulares. Um dos sistemas mais bem estudados para controle genético de formação do padrão é o desenvolvimento embrionário inicial de *Drosophila melanogaster*. Os geneticistas isolaram várias mutações nas moscas-da-fruta que influenciam todos os aspectos de seu desenvolvimento e a análise molecular dessas mutações forneceu muitas informações sobre como os genes controlam o desenvolvimento inicial na *Drosophila*.

Desenvolvimento da mosca-da-fruta

Uma mosca-da-fruta adulta tem três partes básicas do corpo: cabeça, tórax e abdome (**Figura 22.3**). O tórax é composto por três segmentos: o primeiro segmento torácico tem um par de pernas; o segundo, um par de pernas e um par de asas; o terceiro, um par de pernas e os halteres (rudimentos do segundo par de asas encontrados na maioria dos outros insetos). O abdome é composto por vários segmentos.

Quando um ovo de *Drosophila* é fertilizado, seu núcleo diploide (**Figura 22.4 A**) imediatamente se divide nove vezes sem divisão do citoplasma, criando uma única célula multinucleada (**Figura 22.4 B**). Esses núcleos estão dispersos pelo citoplasma, mas posteriormente migram para a periferia do embrião e se dividem mais vezes (**Figura 22.4 C**). A seguir, a membrana celular invagina ao redor de cada núcleo, criando uma camada de aproximadamente 6.000 células na superfície externa do embrião (**Figura 22.4 D**). Os núcleos em uma extremidade do embrião se desenvolvem em células polares, que dão origem às células germinativas. O embrião inicial sofre o desenvolvimento em três estágios distintos: (1) o eixo anteroposterior e o eixo dorsoventral do embrião são estabelecidos (**Figura 22.5 A**); (2) o número e a orientação dos segmentos do corpo são determinados (**Figura 22.5 B**) e (3) a identidade de cada segmento é estabelecida (**Figura 22.5 C**). Diferentes conjuntos de genes controlam cada um desses três estágios (**Quadro 22.1**).

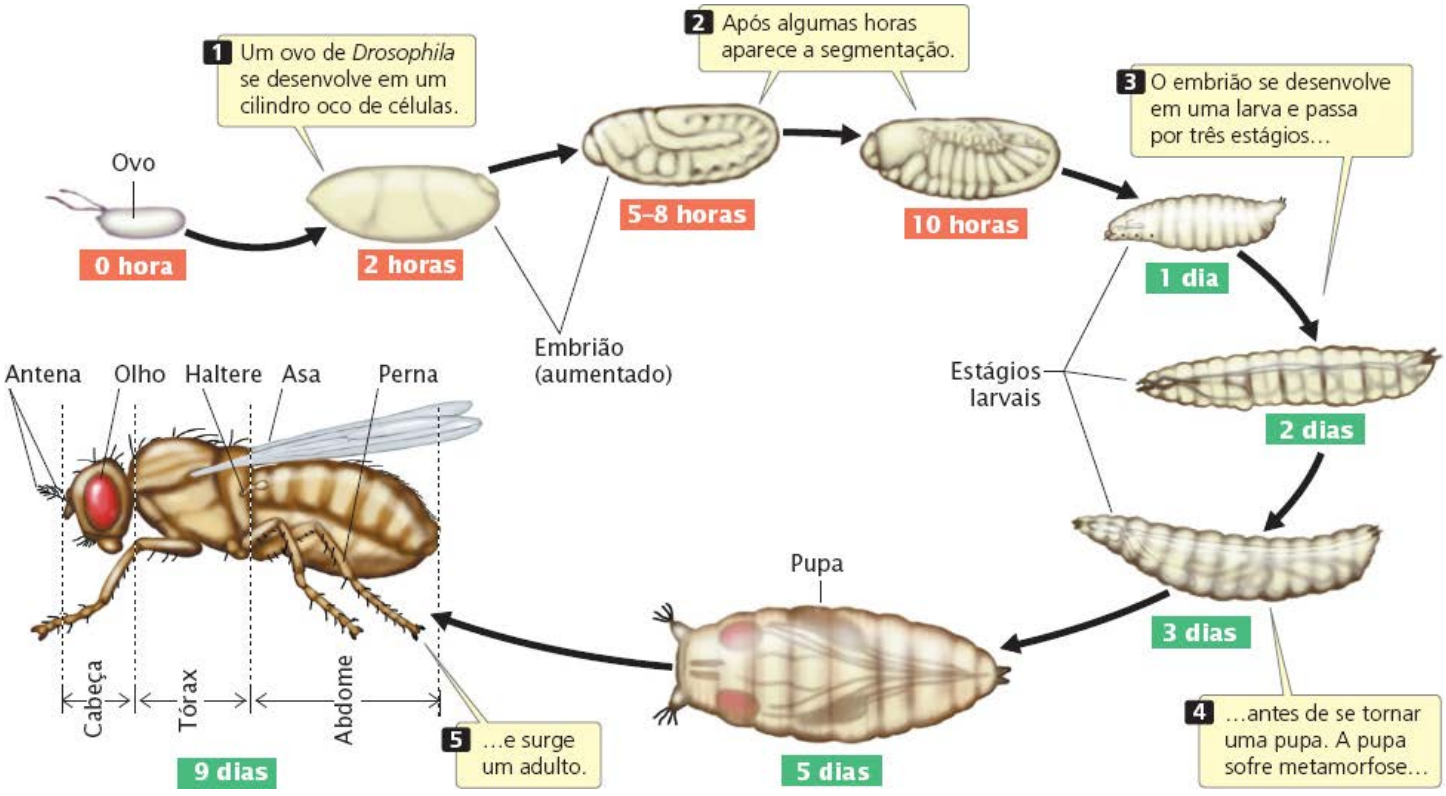


Figura 22.3 A mosca-da-fruta *Drosophila* passa por três estágios larvais e um de pupa antes de se tornar uma mosca adulta. As três principais partes do corpo do adulto são cabeça, tórax e abdome.

Genes da polaridade do ovo

Os **genes da polaridade do ovo** (genes que determinam a polaridade ou direção) são cruciais para estabelecer os dois eixos principais do desenvolvimento nas moscas-da-fruta. Você pode imaginar esses eixos como a longitude e a latitude do desenvolvimento: qualquer localização no embrião da *Drosophila* pode ser definida em relação a esses dois eixos.

Os genes da polaridade do ovo são transcritos em mRNAs durante a formação do ovo na genitora e esses mRNAs são incorporados ao citoplasma do ovo. Após a fertilização, os mRNAs são traduzidos em proteínas importantes para determinar os eixos anteroposterior e dorsoventral do embrião. Como os mRNAs dos genes da polaridade são produzidos pela genitora e influenciam o fenótipo dos descendentes, os traços que eles codificam são exemplos dos efeitos genéticos maternos (ver *Capítulo 5*).

Existem dois conjuntos de genes da polaridade do ovo: um conjunto determina o eixo anteroposterior, e o outro determina o eixo dorsoventral. Esses genes atuam ao ajustar os gradientes de concentração dos morfógenos no embrião em desenvolvimento. Um **morfógeno** é uma proteína com concentração variável e que produz diferentes respostas do desenvolvimento nas diferentes concentrações. Os genes da polaridade do ovo atuam ao produzir proteínas que são distribuídas de forma assimétrica no citoplasma, gerando a polaridade do ovo. Essa distribuição assimétrica pode ocorrer de duas formas. Um mRNA pode estar localizado em regiões específicas do ovo, levando a abundância de proteína nessas regiões onde o mRNA é traduzido. Por outro lado, o mRNA pode ser distribuído de forma aleatória, mas a proteína que ele codifica pode ser distribuída de forma assimétrica por um sistema de transporte que a leva para regiões específicas da célula, pela regulação de sua tradução ou por sua remoção das regiões particulares pela degradação seletiva.

Determinação do eixo dorsoventral. O eixo dorsoventral define as costas (dorso) e o ventre de uma mosca (ver Figura 22.5). Pelo menos 12 diferentes genes determinam esse eixo; um dos mais importantes é um gene chamado *dorsal*. O gene *dorsal* é transcrito e traduzido no ovário materno e o mRNA e a proteína resultantes são transferidos para o ovo durante a ovogênese. Em um ovo recém-posto, o mRNA e a proteína codificados pelo gene *dorsal* estão distribuídos de modo uniforme por todo citoplasma, mas após os núcleos terem migrado para periferia do embrião (ver Figura 22.4 C), a proteína Dorsal é redistribuída. Ao longo de um lado do embrião, a proteína Dorsal permanece no citoplasma; esse lado se tornará a superfície dorsal. Ao longo do outro lado, a proteína Dorsal é recolhida para os núcleos; esse lado se tornará a superfície ventral. Neste ponto, existe um suave gradiente de concentração crescente de proteína Dorsal nuclear a partir do lado dorsal para o ventral (**Figura 22.6**).

Acredita-se que a captação nuclear da proteína Dorsal seja governada por uma proteína chamada Cactus, que se liga à proteína Dorsal e a prende no citoplasma. A presença de outra proteína, chamada Toll, leva à fosforilação da proteína Cactus, fazendo com que ela seja degradada. Quando a proteína Cactus é degradada, a proteína Dorsal é liberada e pode se deslocar para o núcleo. Essas proteínas, Cactus e Toll, regulam a distribuição nuclear da proteína Dorsal, que, por sua vez, determina o eixo dorsoventral do embrião.

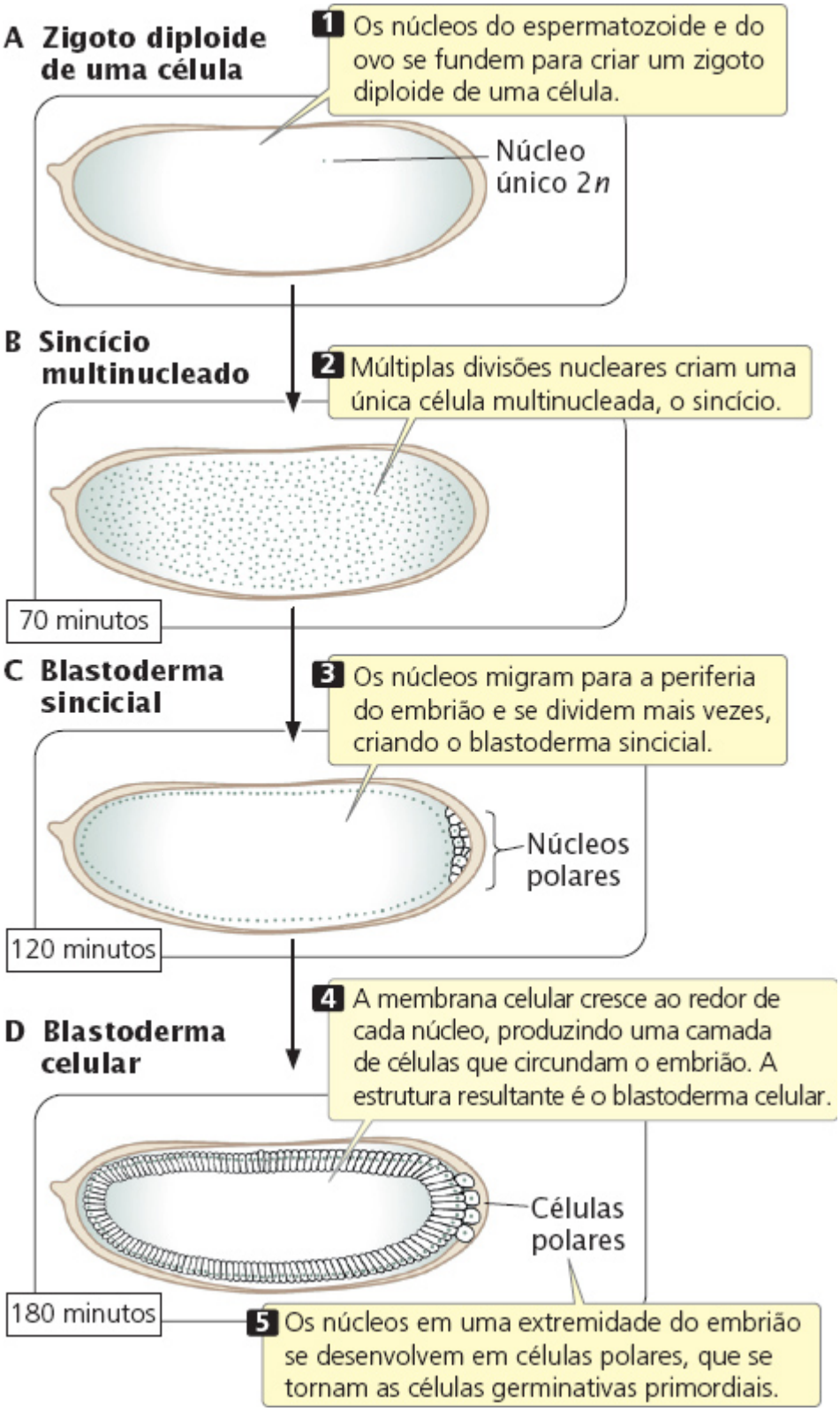


Figura 22.4 Desenvolvimento inicial de um embrião de *Drosophila*.

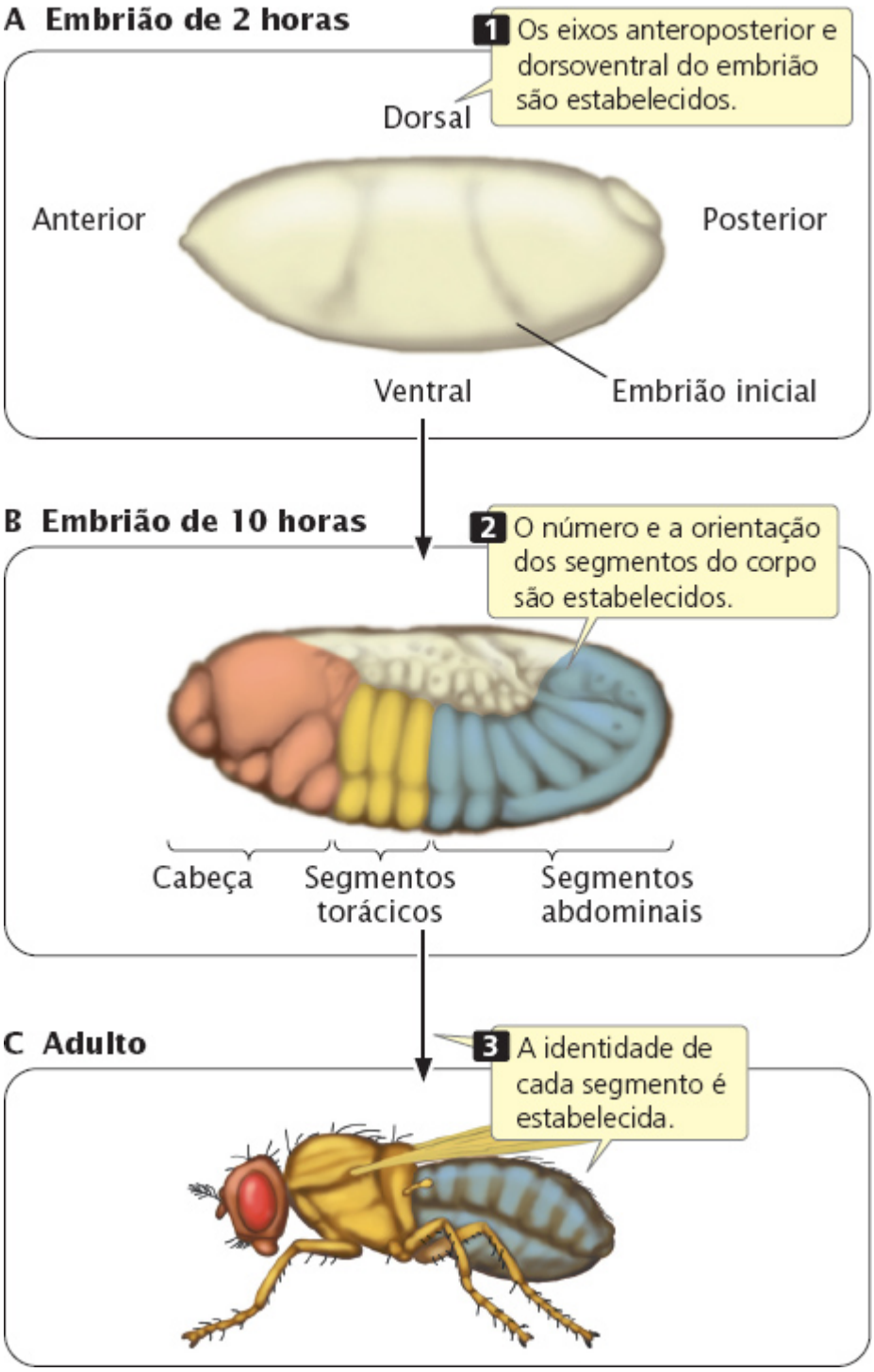


Figura 22.5 **Em um embrião inicial de *Drosophila*, os principais eixos do corpo são estabelecidos, o número e a orientação dos segmentos do corpo são determinados, e a identidade de cada segmento é estabelecida.** Diferentes conjuntos de genes controlam cada um desses três estágios.

Quadro 22.1 Estágios no desenvolvimento inicial das moscas-da-fruta e os genes que controlam cada estágio.

Estágio de desenvolvimento	Genes
Estabelecimento dos principais eixos do corpo	Genes da polaridade do ovo
Determinação do número e polaridade dos segmentos do corpo	Genes de segmentação
Estabelecimento da identidade de cada segmento	Genes homeóticos

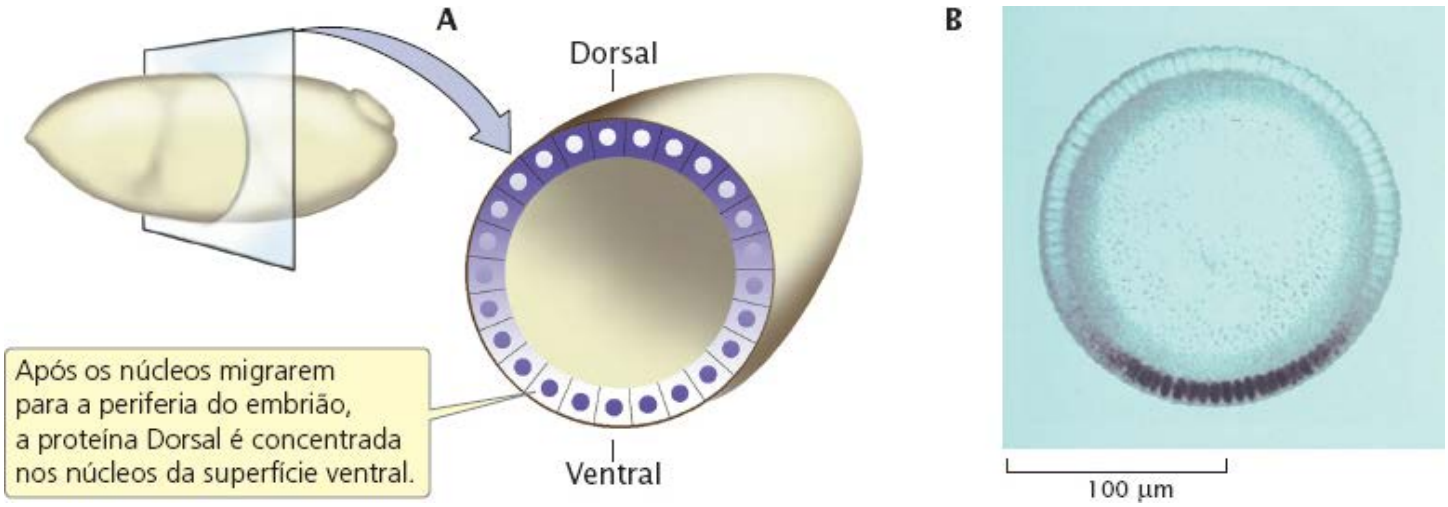


Figura 22.6 A proteína Dorsal nos núcleos ajuda a determinar o eixo dorsalventral do embrião da *Drosophila*. A. concentrações relativas da proteína Dorsal no citoplasma e núcleos das células no embrião inicial da *Drosophila*. **B.** Micrografia de um corte transversal do embrião mostrando a proteína Dorsal, corada escura, nos núcleos ao longo da superfície ventral. (*Parte B:* Max Planck Institute for Developmental Biology.)

No núcleo, a proteína Dorsal atua como fator de transcrição, se ligando a sítios reguladores no DNA e ativando ou reprimindo a expressão de outros genes (**Quadro 22.2**). A elevada concentração nuclear da proteína Dorsal (como nas células no lado ventral do embrião) ativa um gene chamado *twist*, que estimula o desenvolvimento dos tecidos ventrais. As baixas concentrações nucleares da proteína Dorsal (como nas células no lado ventral do embrião) ativam um gene chamado *decapentaplegic*, que estimula o desenvolvimento das estruturas dorsais. Dessa forma, os lados ventral e dorsal do embrião são determinados.

Determinação do eixo anteroposterior. Um dos eventos do desenvolvimento iniciais mais importante é a determinação das extremidades anterior (cabeça) e posterior (traseiro) de um animal. Vamos analisar vários genes importantes que estabelecem esse eixo anteroposterior do embrião da *Drosophila* (**Quadro 22.3**). Um gene importante nesse assunto é o *bicoid*, que é primeiro transcrito no ovário de uma fêmea adulta durante a ovogênese. O mRNA *bicoid* é incorporado ao citoplasma do ovo; à medida que ele passa para o ovo, o mRNA *bicoid* é ancorado à extremidade anterior do ovo em parte de sua extremidade 3′. Essa ancoragem faz com que o mRNA *bicoid* se concentre na extremidade anterior (**Figura 22.7 A**). (Vários outros genes ativos no ovário são necessários para a localização adequada do mRNA *bicoid* no ovo.) Quando o ovo é posto, o mRNA *bicoid* é traduzido na proteína Bicoid. Como a maior parte do mRNA está na extremidade anterior do ovo, a proteína Bicoid é sintetizada nessa região e forma um gradiente de concentração ao longo do eixo anteroposterior do embrião, com uma alta concentração na extremidade anterior e uma baixa concentração na extremidade posterior. Esse gradiente é mantido pela síntese contínua da proteína Bicoid e sua meia-vida curta.

A elevada concentração da proteína Bicoid na extremidade anterior induz o desenvolvimento das estruturas anteriores como a cabeça da mosca-da-fruta. Ela estimula o desenvolvimento das estruturas anteriores ao se ligar a sequências reguladoras no DNA e influenciar a expressão de outros genes. Um dos genes mais importantes estimulados pela proteína Bicoid é o *hunchback*, necessário para o desenvolvimento da cabeça e das estruturas torácicas da mosca-da-fruta.

Quadro 22.2 Genes importantes que controlam o desenvolvimento do eixo dorsoventral nas moscas-da-fruta e sua ação.

Gene	Local em que é expresso	Ação do produto do gene
dorsal	Ovário	Afeta a expressão de genes como o *twist* e *decapentaplegic*
cactus	Ovário	Prende a proteína Dorsal no citoplasma
toll	Ovário	Leva à fosforilação da proteína Cactus, que é então degradada, liberando a proteína Dorsal para se deslocar para o núcleo das células ventrais
twist	Embrião	Participa do desenvolvimento dos tecidos

		mesodermais*
decapentaplegic	Embrião	Participa do desenvolvimento das estruturas do intestino

*Uma das três camadas principais de tecido no embrião inicial.

Quadro 22.3 Alguns genes importantes que determinam o eixo anteroposterior nas moscas-da-fruta.

Gene	Local em que é expresso	Ação
bicoid	Ovário	Regula a expressão de genes responsáveis pelas estruturas anteriores, estimula o *hunchback*
nanos	Ovário	Regula a expressão de genes responsáveis pelas estruturas posteriores, inibe a tradução do mRNA *hunchback*
hunchback	Embrião	Regula a transcrição de genes responsáveis pelas estruturas anteriores

O desenvolvimento do eixo anteroposterior também é muito influenciado por um gene chamado *nanos*, um gene da polaridade do ovo que atua na extremidade posterior do eixo. O gene *nanos* é transcrito na fêmea adulta e o mRNA resultante está localizado na extremidade posterior do ovo (**Figura 22.7 B**). Após a fertilização, o mRNA *nanos* é traduzido na proteína Nanos, que se difunde lentamente na direção da extremidade anterior. O gradiente da proteína Nanos é oposto ao da proteína Bicoid: Nanos está mais concentrada na extremidade posterior do embrião e está menos concentrada na extremidade anterior. A proteína Nanos inibe a formação de estruturas anteriores ao reprimir a tradução do mRNA *hunchback*. A síntese da proteína Hunchback é estimulada na extremidade anterior do embrião pela proteína Bicoid e é reprimida na extremidade posterior pela proteína Nanos. Esse estímulo e repressão combinados geram um gradiente de concentração da proteína Hunchback ao longo do eixo anteroposterior, que, por sua vez, afeta a expressão de outros genes ao ajudar a determinar as estruturas anteriores e posteriores.

Conceitos

Os principais eixos do desenvolvimento nos embriões iniciais da mosca-da-fruta são estabelecidos como resultado de diferenças iniciais na distribuição de mRNAs específicos e proteínas codificados por genes na genitora (efeito genético materno). Essas diferenças na distribuição estabelecem gradientes de concentração de morfógenos, o que faz com que diferentes genes sejam ativados em diferentes partes do embrião.

Checagem dos conceitos 2

A elevada concentração de qual proteína estimula o desenvolvimento das estruturas anteriores?

a. Dorsal.
b. Toll.
c. Bicoid.
d. Nanos.

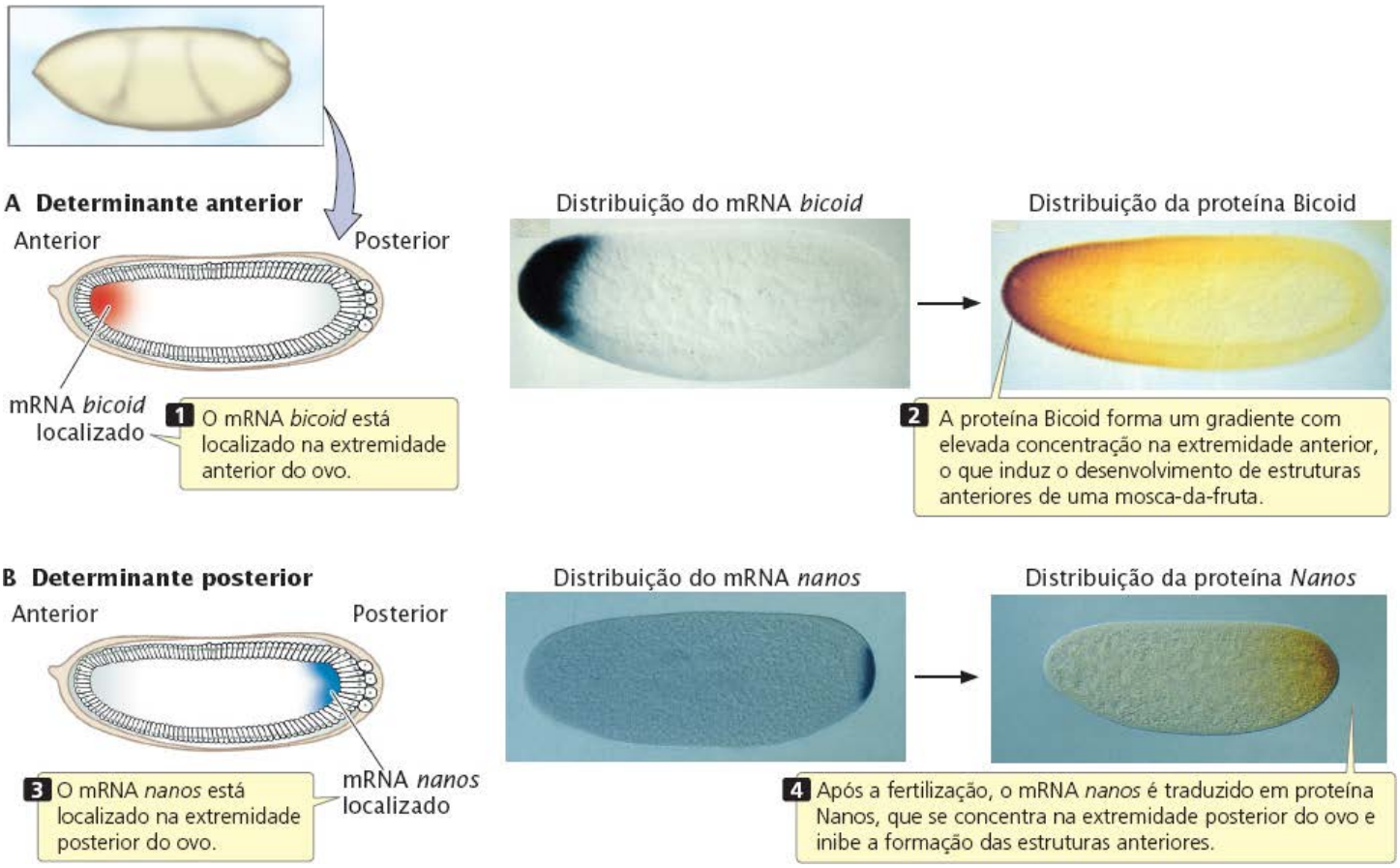


Figura 22.7 O eixo anteroposterior em um embrião de *Drosophila* é determinado pelas concentrações das proteínas Bicoid e Nanos. (*Partes A e B:* De Christiane Nusslein-Volhard, "Determination of the Embryonic Axes of Drosophilia," Development, Suppl. 1, 1991, 1. © Company of Biologists. *Partes C e D:* Cortesia de E. R. Gavis, L. K. Dickenson, and R. Lehman, Massachusetts Institute of Technology.)

Genes de segmentação

Como todos os insetos, a mosca-da-fruta tem um plano de corpo segmentado. Quando os eixos básicos dorsoventral e anteroposterior do embrião da mosca-da-fruta estão estabelecidos, os **genes de segmentação** controlam a diferenciação do embrião em segmentos individuais. Esses genes afetam o número e a organização dos segmentos e as mutações nesses genes em geral rompem conjuntos inteiros de segmentos. Os aproximadamente 25 genes de segmentação na *Drosophila* são transcritos após a fertilização, então eles não exibem um efeito genético materno e sua expressão é regulada pelos gradientes das proteínas Bicoid e Nanos.

Os genes de segmentação são divididos em três classes como apresentado no **Quadro 22.4** e na **Figura 22.8**. As três classes atuam em sequência, afetando de forma progressiva as regiões menores do embrião. Primeiro, os produtos dos genes da polaridade do ovo ativam ou reprimem os **genes *gap***, que dividem o embrião em grandes regiões. Os genes *gap*, por sua vez, regulam os **genes da regra dos pares**, que afetam o desenvolvimento dos pares de segmentos. Finalmente, os genes da regra dos pares influenciam os **genes de polaridade segmentar**, que guiam o desenvolvimento dos segmentos individuais.

Quadro 22.4 Genes de segmentação e os efeitos das suas mutações.

Classe do gene	Efeitos das mutações	Exemplos de genes
Genes *gap*	Eliminam segmentos adjacentes	*hunchback, Krüppel, knirps, giant, tailless*
Genes da regra dos pares	Eliminam a mesma parte do padrão em segmento alternado	*runt, hairy, fushi tarazu, even-skipped, odd-paired, sloppy paired, odd-skipped*
Genes de polaridade segmentar	Afetam a polaridade segmentar, parte do	*engrailed, wingless, gooseberry, cubitus interruptus, patched, hedgehog, disheveled, costal-2, fused*

segmento substituído
pela imagem de espelho
de parte de outro
segmento

Os genes *gap* definem grandes seções do embrião; as mutações nesses genes eliminam grupos inteiros de segmentos adjacentes. As mutações no gene *Krüppel*, por exemplo, levam à ausência de vários segmentos adjacentes. Os genes da regra dos pares definem seções regionais do embrião e afetam os segmentos alternados. Mutações no gene *even-skipped* provocam a deleção de segmentos de números pares enquanto as mutações no gene *fushi tarazu* provocam a ausência de segmentos ímpares. Os genes de polaridade segmentar afetam a orientação dos segmentos. Mutações nesses genes fazem com que parte de cada segmento seja eliminada e substituída por uma imagem em espelho de parte ou de todo um segmento adjacente. Por exemplo, mutações no gene *gooseberry* fazem com que a metade posterior de cada segmento seja substituída pela metade anterior de um segmento adjacente.

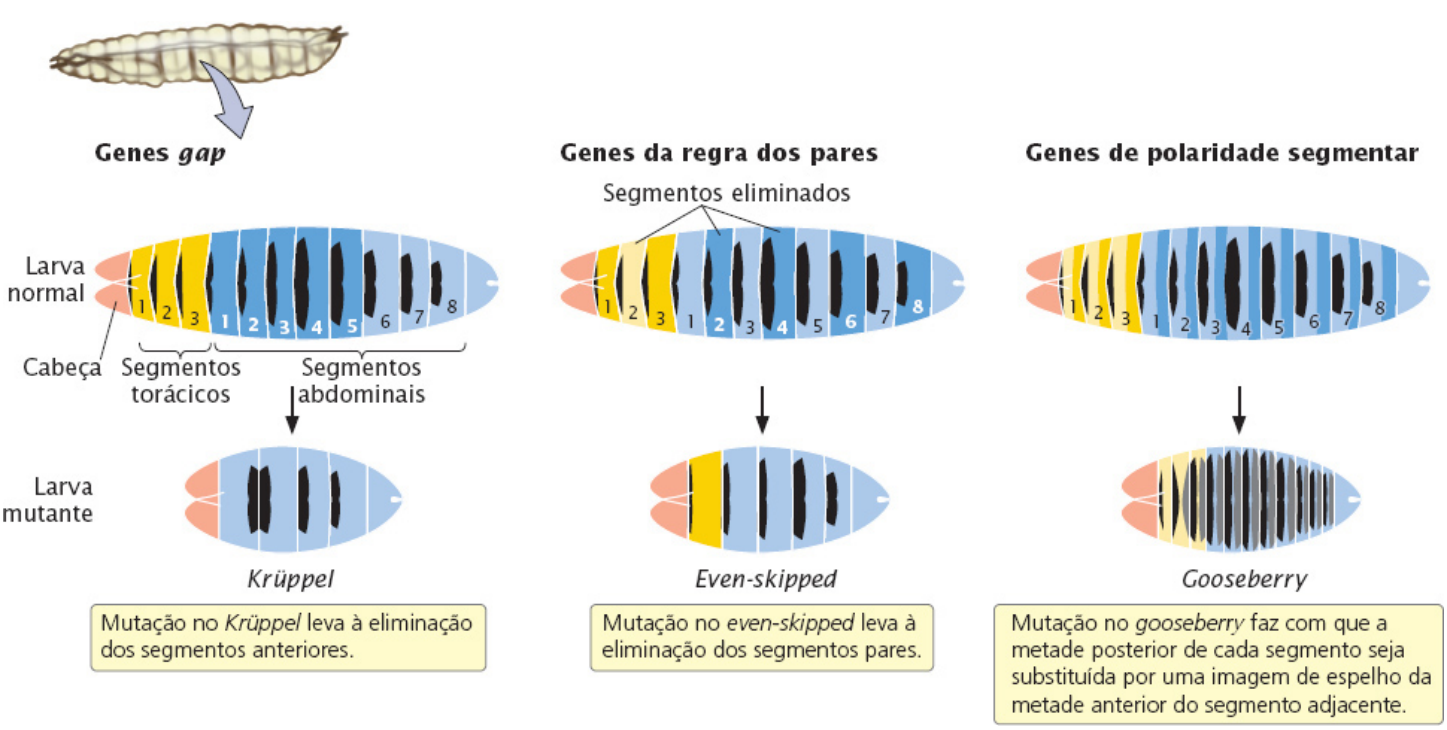


Figura 22.8 Os genes de segmentação controlam a diferenciação do embrião de *Drosophila* em segmentos individuais. Os genes *gap* afetam grandes secções do embrião. Os genes da regra dos pares afetam segmentos alternados. Os genes de polaridade segmentar afetam a orientação dos segmentos.

Conceitos

Quando os principais eixos do embrião da mosca-da-fruta são estabelecidos, os genes da segmentação determinam o número, a orientação e a organização básica dos segmentos do corpo.

Checagem dos conceitos 3

A sequência correta na qual os genes de segmentação atuam é:

a. Genes de polaridade segmentar → genes *gap* → genes da regra dos pares
b. Genes *gap* → genes da regra dos pares → genes de polaridade segmentar
c. Genes de polaridade segmentar → genes da regra dos pares → genes *gap*
d. Genes *gap* → genes de polaridade segmentar → genes da regra dos pares

Genes homeóticos em *Drosophila*

Após os genes de segmentação terem estabelecido o número e a orientação dos segmentos, os **genes homeóticos** são ativados e determinam a *identidade* dos segmentos individuais. Os olhos surgem normalmente apenas no segmento da cabeça, enquanto as pernas se desenvolvem apenas nos segmentos torácicos. Os produtos dos genes homeóticos ativam outros genes que codificam essas características específicas de segmento. As mutações nos genes homeóticos fazem com que partes do corpo surjam nos

segmentos errados.

No final dos anos 1940, Edward Lewis começou a estudar as mutações homeóticas em *Drosophila*, mutações que provocam rearranjos bizarros das partes do corpo. As mutações no gene *Antennapedia*, por exemplo, fazem com que as pernas se desenvolvam na cabeça de uma mosca no lugar das antenas. Os genes homeóticos criam locais para as células de segmentos específicos, dizendo às células onde elas estão nas regiões definidas pelos genes de segmentação. Quando um gene homeótico sofre uma mutação, o "endereço" está errado e as células no segmento se desenvolvem como se estivessem em outra parte no embrião.

Os genes homeóticos em *Drosophila* são expressos após a fertilização e ativados por concentrações específicas de proteínas produzidas pelos genes *gap*, da regra dos pares e de polaridade segmentar. O gene homeótico *Ultrabithorax* (*Ubx*), por exemplo, é ativado quando a concentração da proteína Hunchback (um produto de um gene *gap*) está dentro de alguns valores. Essas concentrações existem apenas na região média do embrião, então o *Ubx* é expresso apenas nesses segmentos.

Os genes homeóticos nos animais codificam proteínas reguladoras que se ligam ao DNA; cada gene tem um subconjunto de nucleotídios, chamado de ***homeobox***, semelhante em todos os genes homeóticos. O *homeobox* codifica 60 aminoácidos que servem como um domínio de ligação ao DNA, esse domínio está relacionado com o motivo hélice-giro-hélice (ver Figura 16.2 A). Os *homeoboxes* também são encontrados nos genes de segmentação e em outros genes importantes para o desenvolvimento regional.

Existem dois grupos principais dos genes homeóticos em *Drosophila*. Um grupo, o **complexo *Antennapedia***, afeta o desenvolvimento da cabeça e dos segmentos torácicos anteriores da mosca adulta. O outro grupo é composto pelo **complexo bitórax** e inclui os genes que influenciam os segmentos torácico posterior e abdominal da mosca adulta. Juntos, os genes *bithorax* e *Antennapedia* são chamados de **complexo homeótico** (HOM-C). Em *Drosophila*, o complexo bitórax inclui três genes e o complexo *Antennapedia* tem cinco, todos estão localizados no mesmo cromossomo (**Figura 22.9**). Além desses oito genes, o HOM-C tem muitas sequências que regulam os genes homeóticos. Assim, a ordem dos genes no HOM-C é a mesma que a ordem na qual os genes são expressos ao longo do eixo anteroposterior do corpo. Os genes que são expressos nos segmentos mais anteriores são encontrados em uma extremidade do complexo, enquanto os genes expressos na extremidade mais posterior do embrião são encontrados na outra extremidade do complexo (ver Figura 22.9).

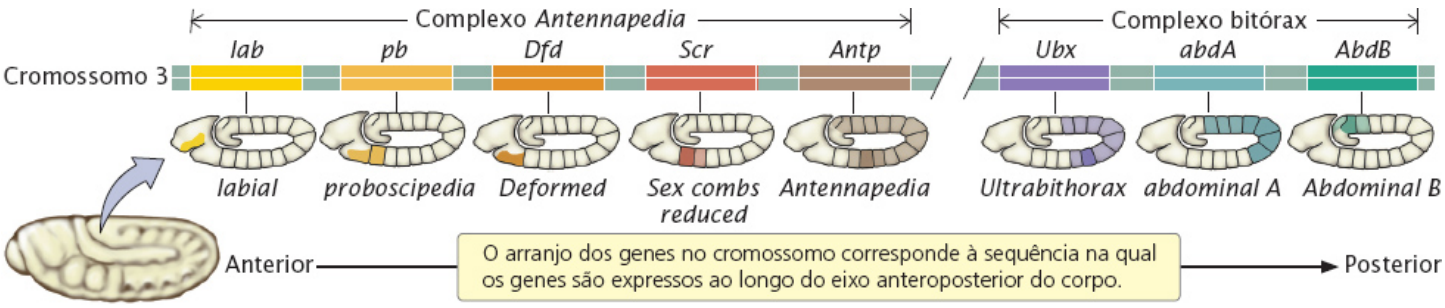


Figura 22.9 Os genes homeóticos, que determinam a identidade dos segmentos individuais em *Drosophila*, estão presentes em dois complexos. O complexo *Antennapedia* tem cinco genes, o complexo bitórax tem três genes.

Conceitos

Os genes homeóticos ajudam a determinar a identidade de segmentos individuais nos embriões de *Drosophila* ao produzir proteínas que se ligam ao DNA que ativam outros genes. Cada gene homeótico tem uma sequência consenso chamada *homeobox*, que codifica o domínio de ligação do DNA.

Checagem dos conceitos 4

As mutações nos genes homeóticos causam:

a. A deleção nos segmentos.
b. A ausência das estruturas.
c. Segmentos em excesso.
d. As estruturas surgirem no lugar errado.

Genes *homeobox* em outros organismos

Após os genes homeóticos em *Drosophila* terem sido isolados e clonados, os geneticistas moleculares resolveram determinar se existem genes semelhantes em outros animais; foram usadas sondas complementares a *homeobox* de *Drosophila* para pesquisa

de genes homólogos que possam participar do desenvolvimento de outros animais. A pesquisa foi um grande sucesso: foram encontrados genes com *homeobox* em todos os animais, incluindo nematódeos, escaravelhos, ouriços-do-mar, sapos, pássaros e mamíferos. Foram descobertos genes com *homeoboxes* até em fungos e plantas, indicando que *homeobox* surgiu no início da evolução dos eucariotos. Um grupo de genes *homeobox* inclui os **genes *Hox***, que incluem o complexo homeótico de *Drosophila* descrito na seção anterior. Os genes *Hox* são encontrados em todos os animais, exceto nas esponjas.

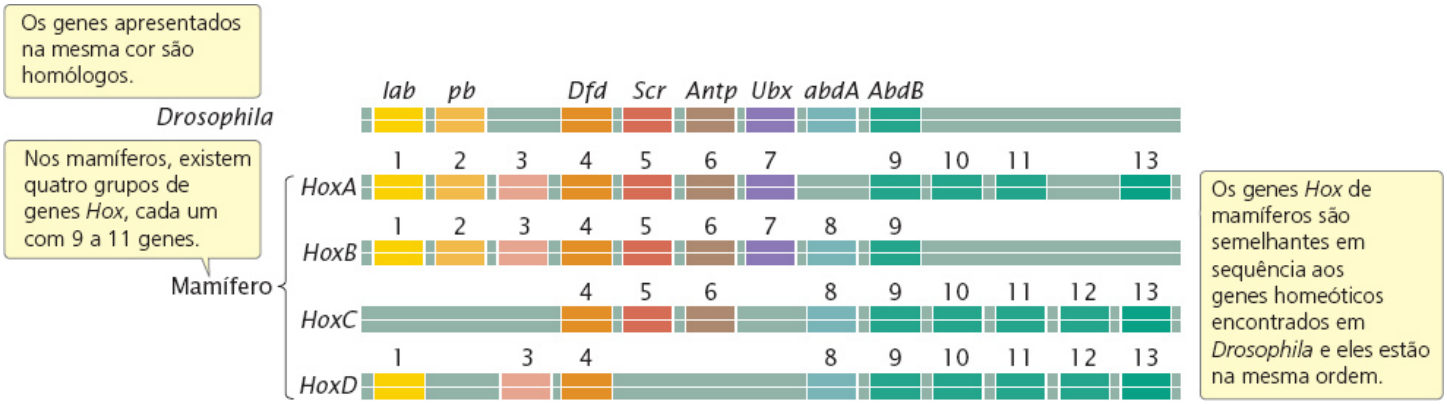


Figura 22.10 Os genes *Hox* nos mamíferos são semelhantes aos encontrados em *Drosophila*. Os complexos são organizados de modo que os genes com sequências semelhantes fiquem na mesma coluna. Consulte a Figura 22.9 sobre os nomes completos dos genes de *Drosophila*.

Nos vertebrados, existem quatro grupos de genes *Hox*, cada um com 9 a 11 genes. Os genes *Hox* nos mamíferos, como os de *Drosophila*, codificam fatores de transcrição que ajudam a determinar a identidade das regiões do corpo ao longo do eixo anteroposterior. Os genes *Hox* de outros organismos exibem a mesma relação entre a ordem no cromossomo e a ordem de sua expressão ao longo do eixo anteroposterior do embrião como em *Drosophila* (**Figura 22.10**), mas esse padrão não é universal. Por exemplo, o tunicado *Oikopleura dioica* (um parente primitivo dos vertebrados) tem 9 genes *Hox*, mas os genes estão dispersos pelo genoma, diferente do arranjo agrupado observado na maioria dos animais. Apesar da falta de agrupamento físico, os genes *Hox* de *O. dioica* são expressos na mesma ordem anteroposterior observada nos vertebrados.

Os genes *Hox* dos vertebrados também exibem uma relação entre sua ordem no cromossomo e no momento de sua expressão: os genes em uma extremidade do complexo (os expressos na extremidade anterior) são expressos no início do desenvolvimento, enquanto os genes na outra extremidade (os expressos na extremidade posterior) são expressos posteriormente. Se um gene *Hox* é deslocado experimentalmente para uma nova localização no complexo de gene *Hox*, ele é expresso no tecido adequado, mas o momento de sua expressão é alterado, sugerindo que o momento da expressão do gene é controlado por sua localização física no complexo. Embora o mecanismo de seu controle sequencial não esteja bem compreendido, evidências sugerem que, nos camundongos, ele é controlado por uma mudança progressiva nos padrões de metilação das proteínas histona, uma mudança epigenética que altera a estrutura da cromatina e afeta a transcrição (ver *Capítulo 21*). Estudos recentes também identificaram genes de microRNAs (ver *Capítulo 14*) nos grupos do gene *Hox*, e evidências sugerem que miRNAs são importantes para controlar a expressão de alguns desses genes.

Os genes *Hox* e sua expressão estão relacionados com mudanças anatômicas nos animais, e existe uma hipótese de que esses genes sejam importantes na evolução dos animais. Por exemplo, o anfioxo *Branchiostoma* (outro parente primitivo dos vertebrados) tem uma forma de corpo simples e apenas 14 genes *Hox* em um único grupo, enquanto os peixes com uma arquitetura corpórea muito mais complexa têm 48 genes *Hox* em sete grupos.

Conceitos conectantes

Controle do desenvolvimento

O desenvolvimento é um processo complexo composto por numerosos eventos que devem ocorrer em uma sequência muito específica. Estudos com as moscas-da-fruta e outros organismos revelam que esse processo é regulado por vários genes. Em *Drosophila*, o eixo dorsoventral e o eixo anteroposterior são estabelecidos pelos genes maternos; esses genes codificam mRNAs e proteínas que ficam localizados em regiões específicas no ovo e fazem com que genes específicos sejam expressos em diferentes regiões do embrião. As proteínas desses genes então estimulam outros genes, que por sua vez estimulam outros genes em uma cascata de controle. Como seria esperado, a maioria dos produtos dos genes na cascata são proteínas reguladoras, que se ligam ao DNA e ativam outros genes.

No curso do desenvolvimento, regiões sucessivamente menores do embrião são determinadas (**Figura**

22.11). Em *Drosophila*, primeiro, os principais eixos e regiões do embrião são estabelecidos pelos genes da polaridade do ovo. A seguir, os padrões em cada região são determinados pela ação dos genes de segmentação: os genes *gap* definem as seções maiores, os genes da regra dos pares definem as seções regionais do embrião e afetam os segmentos alternados, e os genes de polaridade segmentar afetam os segmentos individuais. Finalmente, os genes homeóticos fornecem uma identidade única a cada segmento. Os gradientes iniciais nas proteínas e mRNA estimulam a expressão gênica localizada, que produz gradientes mais especificamente localizados que estimulam a expressão gênica mais localizada ainda. A regulação do desenvolvimento se torna cada vez mais sutil.

Os processos pelos quais os membros, órgãos e tecidos se formam (chamados de morfogênese) ainda são pouco compreendidos, embora esse padrão de expressão gênica generalizada para expressão gênica localizada seja encontrado com frequência. **Resolva o Problema 22**

Mudanças epigenéticas no desenvolvimento

O desenvolvimento inicial da mosca-da-fruta é controlado em grande parte pelos produtos de alguns genes-chave que ativam ou reprimem de forma seletiva a expressão de outros genes. Como foi discutido no Capítulo 21, a expressão gênica nos eucariotos é afetada por mudanças epigenéticas e de fato, a epigenética é importante para o desenvolvimento.

No Capítulo 21, as mudanças epigenéticas foram definidas como alterações hereditárias na estrutura do DNA e da cromatina, alterações que afetam a expressão gênica e são transmitidas para outras células, mas não são mudanças na sequência de bases do DNA. No curso do desenvolvimento, um zigoto de uma célula se divide e dá origem a muitas células, que se diferenciam e adquirem as características de órgãos e tecidos específicos. Cada tipo de célula expressa um subconjunto diferente de genes, produzindo as proteínas necessárias para aquele tipo de célula, e esse programa de expressão gênica é transmitido quando a célula diferenciada se divide.

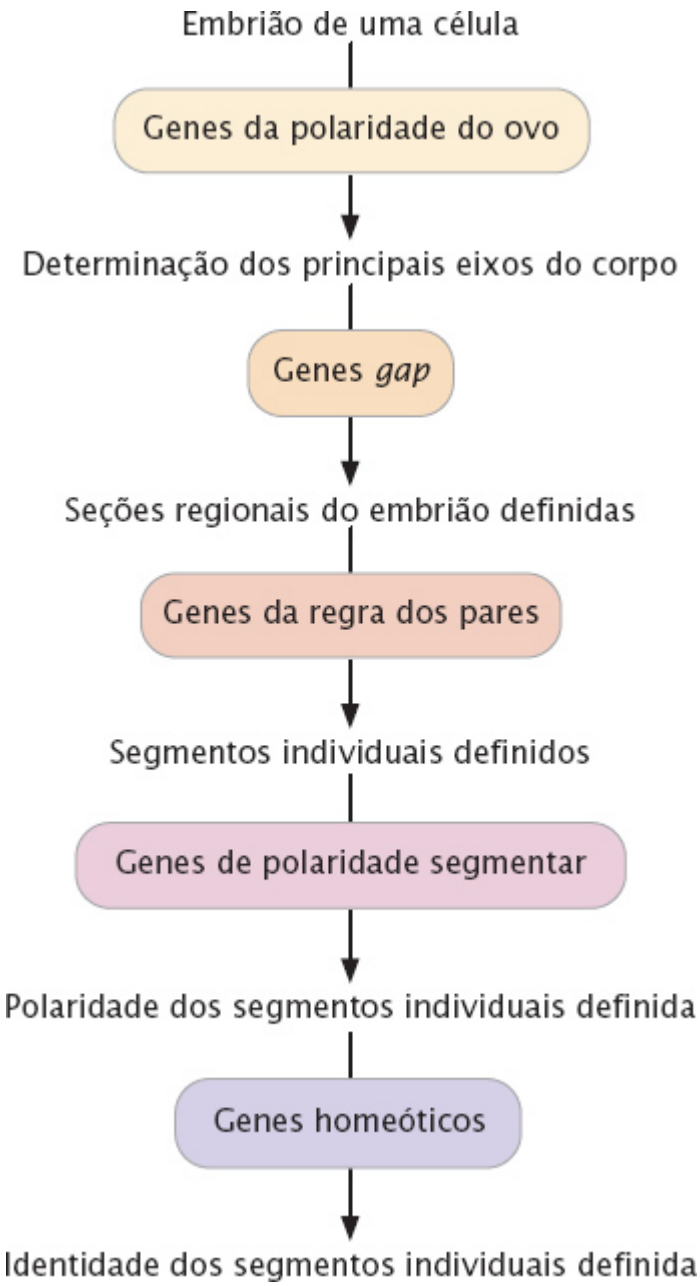


Figura 22.11 Uma cascata da regulação gênica determina a polaridade e a identidade de segmentos individuais de *Drosophila*. No curso do desenvolvimento, são determinadas regiões sucessivamente menores do embrião.

O programa de expressão gênica das células que compõem um determinado órgão ou tipo de tecido é definido pelas marcas epigenéticas. À medida que o desenvolvimento e a diferenciação seguem, as células adquirem mudanças epigenéticas que ligam e desligam conjuntos específicos de genes. No Capítulo 21, analisamos vários tipos de marcas epigenéticas, incluindo a metilação do DNA, a modificação das proteínas histona e mudanças causadas por moléculas de RNA. Essas mudanças epigenéticas ajudam a determinar quais genes são expressos por uma célula. Nos estágios iniciais do desenvolvimento, os genes que podem ser necessários nos estágios posteriores do desenvolvimento são mantidos em estado de silêncio temporário pelas modificações de histona. As modificações da histona são flexíveis e revertidas com facilidade e esses genes podem ser ativados posteriormente nos estágios do desenvolvimento. Durante o desenvolvimento, outros genes são silenciados de forma permanente; esse silenciamento mais duradouro é obtido com a metilação do DNA.

22.3 Os genes controlam o desenvolvimento das flores nas plantas

Examinamos em detalhes a formação padrão em *Drosophila*, que serve como um sistema modelo de desenvolvimento. Outro sistema modelo que fornece discernimento importante em como os genes influenciam os padrões de crescimento e desenvolvimento é a formação das partes das flores nos angiospermas.

Um dos eventos do desenvolvimento mais importantes na vida de uma planta é a troca do crescimento vegetativo para a floração. O momento preciso dessa troca é afetado pela estação, duração do dia, tamanho da planta e vários outros fatores e está sob controle de vários genes diferentes. O desenvolvimento da flor por si só também está sob controle genético e os genes

homeóticos são cruciais na determinação das estruturas florais.

Anatomia da flor

Uma flor é composta de quatro anéis concêntricos de folhas modificadas, chamadas verticilos. O verticilo mais externo (verticilo 1) é composto por sépalas verdes semelhantes a folhas. O verticilo seguinte (verticilo 2) é composto de pétalas, que em geral não têm clorofila. O verticilo 3 é composto de estames, que carreiam o pólen, e o verticilo 4 é composto por carpelos que são fundidos para formar o estigma que carreia os ovos. Na *Arabidopsis* do tipo selvagem, uma planta modelo genético (consulte o *Guia de Referência para Organismos-Modelo da Genética*), existem quatro sépalas, quatro pétalas brancas, seis estames (quatro longos e dois curtos) e dois carpelos (**Figura 22.12 A**).

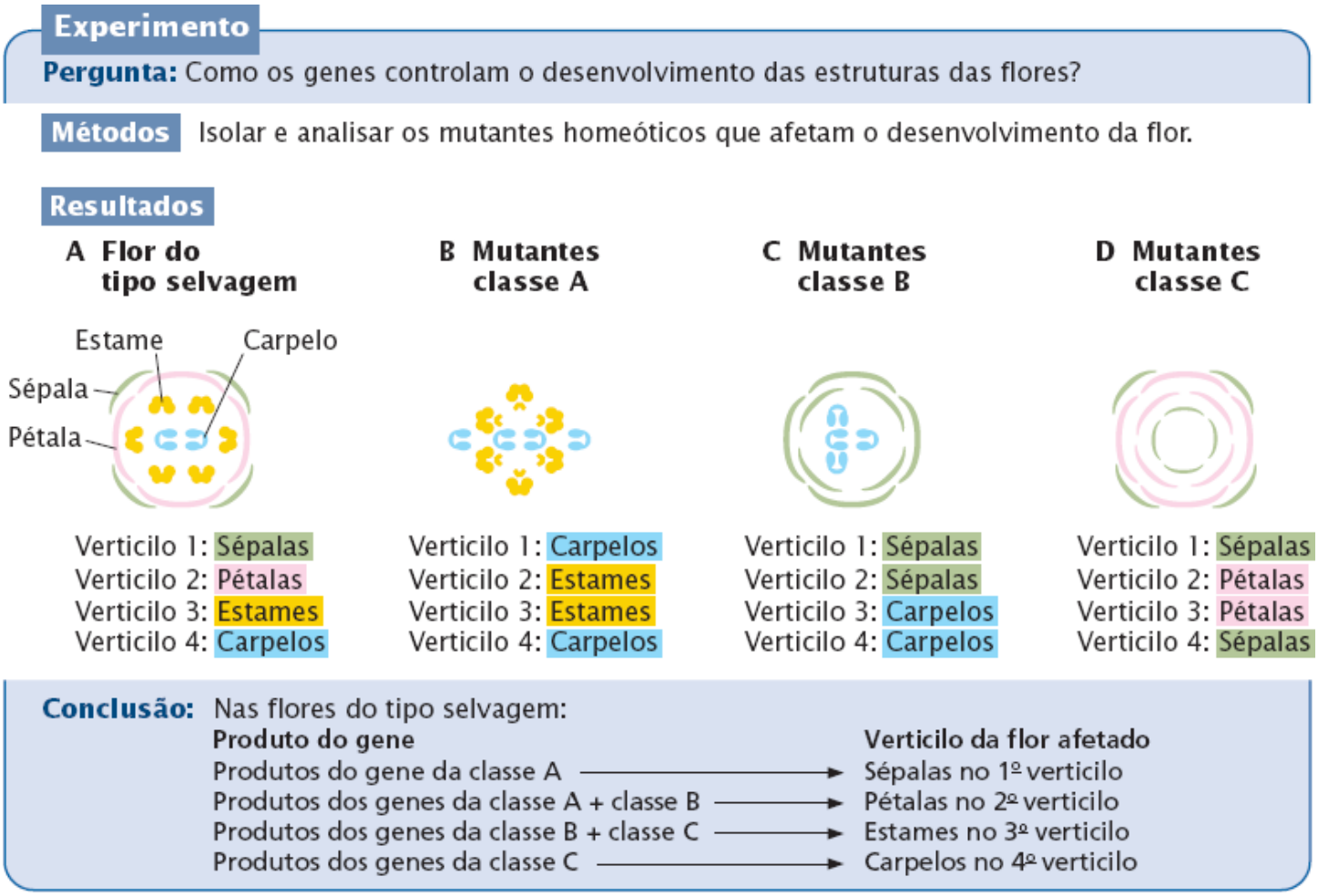


Figura 22.12 A análise dos mutantes homeóticos em *Arabidopsis thaliana* permitiu a compreensão dos genes que determinam as estruturas florais nas plantas.

Controle genético de desenvolvimento da flor

Elliot Meyerowitz e colaboradores fizeram uma série de experimentos para examinar a base genética do desenvolvimento da flor em *Arabidopsis*. Eles começaram isolando e analisando as mutações homeóticas em *Arabidopsis*. As mutações homeóticas foram identificadas pela primeira vez nas plantas em 1894, quando William Bateson observou que algumas partes florais das plantas surgiam no lugar errado: ele descobriu, por exemplo, flores nas quais os estames cresceram no lugar onde as pétalas crescem.

Meyerowitz e colaboradores usaram esses tipos de mutantes para revelar os genes que controlam o desenvolvimento da flor. Eles conseguiram colocar as mutações homeóticas que isolaram em três grupos com base no seu efeito sobre a estrutura floral. Os mutantes classe A tinham carpelos em vez de sépalas no primeiro verticilo e estames em vez de pétalas no segundo verticilo (**Figura 22.12 B**). O terceiro verticilo tinha estames e o quarto tinha carpelos, o padrão normal. Os mutantes classe B tinham sépalas no primeiro e no segundo verticilos e carpelos nos terceiro e quarto verticilos (**Figura 22.12 C**). O grupo final, mutantes classe C, tinha sépalas e pétalas no primeiro e segundo verticilos, respectivamente, como o normal, mas tinha pétalas no terceiro verticilo e sépalas no quarto (**Figura 22.12 D**).

Meyerowitz e colaboradores concluíram que cada classe de mutantes não tinha o produto de um gene ou os produtos de um conjunto de genes críticos para o desenvolvimento adequado da flor: os mutantes classe A não apresentavam a atividade do

gene A, os mutantes classe B não apresentavam atividade do gene B, e os mutantes da classe C não apresentavam atividade do gene C. Eles criaram uma hipótese de que os genes A eram ativos no primeiro e segundo verticilos. Os produtos do gene da classe A sozinhos faziam com que o primeiro verticilo se diferenciasse em sépalas, e junto com os produtos do gene da classe B eles faziam com que o segundo verticilo se desenvolvesse em pétalas. Os produtos do gene da classe C junto com os produtos dos genes da classe B induziam o terceiro verticilo a se desenvolver em estames. Os produtos do gene da classe C faziam com que o quarto verticilo se tornasse carpelos. Os produtos de diferentes classes de genes e seus efeitos estão resumidos na conclusão da Figura 22.12.

Para explicar os resultados, eles também propuseram que os genes de algumas classes afetam a atividades de outros. Onde a classe A está ativa, a classe C é reprimida, e onde a classe C está ativa, a classe A é reprimida. Além disso, se uma mutação inativa A, então C se torna ativo e vice-versa. Os genes da classe A são normalmente expressos nos verticilos 1 e 2, os genes das classes B são expressos nos verticilos 2 e 3 e os genes da classe C são expressos nos verticilos 3 e 4 (**Figura 22.13**).

A interação dessas três classes de genes explica as diferentes classes de mutantes na Figura 22.12. Por exemplo, os mutantes da classe A não têm os produtos dos genes da classe A e, portanto, os genes da classe C estão ativos em todos os tecidos porque, quando A está inativo, C se torna ativo. Portanto, o verticilo 1, com apenas os produtos do gene da classe C, terá carpelos; o verticilo 2, com produtos dos genes das classes C e B, produzirá estames; o verticilo 3, com produtos dos genes das classes B e C, produzirá estames; e o verticilo 4, com apenas atividade do gene da classe C, produzirá carpelos (ver Figura 22.12 B):

Produtos do gene classe C (na ausência de classe A)	→ Carpelos (1º verticilo)
Produtos do gene classe B + classe C (na ausência de classe A)	→ Estames (2º verticilo)
Produtos dos genes de classe B + classe C	→ Estames (3º verticilo)
Produtos do gene da classe C	→ Carpelos (4º verticilo)

Para confirmar essa explicação, Meyerowitz e colaboradores cruzaram mutantes duplos e triplos e previram o desfecho. As estruturas florais resultantes se encaixavam nas suas previsões. Em estudos posteriores, eles isolaram os genes de cada classe. Existem dois genes da classe A, chamados *APETALA1* (*AP1*) e *APETALA2* (*AP2*); dois genes da classe B, chamados *APETALA3* (*AP3*) e *PISTILLATA* (*PI*), e um gene da classe C chamado *AGAMOUS* (*AG*). A clonagem e o sequenciamento desses genes revelaram que todos são genes *MADS-box*, genes que atuam como fatores de transcrição e afetam a expressão de outros genes. Os genes *MADS-box* nas plantas têm um papel semelhante ao dos genes *homeobox* nos animais, embora os genes *MADS-box* e *homeobox* não sejam homólogos.

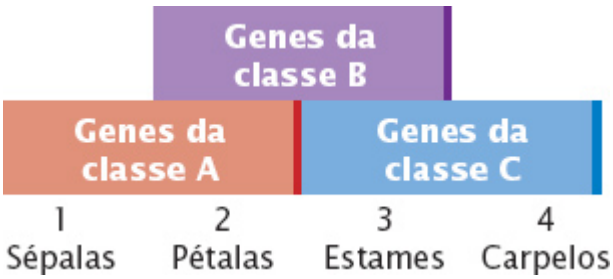


Figura 22.13 A expressão dos genes das classes A, B e C varia entre as estruturas de uma flor.

Os resultados de outros estudos demonstraram a presença de um grupo adicional de genes chamados *SEPALLATA* (*SEP*) que são expressos nos verticilos 2, 3 e 4 e também são necessários para o desenvolvimento floral normal. Se os genes *SEP* estiverem defeituosos, a flor será composta inteiramente de sépalas. Os achados a partir de estudos de outras espécies demonstraram que esse sistema de desenvolvimento de planta não existe apenas em *Arabidopsis*, mas também em outras plantas florescentes. É importante observar que esses genes são necessários, mas não suficientes para o desenvolvimento adequado da flor; outros genes também participam na identidade de partes diferentes das flores. Resolva o Problema 25

Conceitos

Os genes homeóticos das plantas controlam o desenvolvimento de suas estruturas florais. Os produtos de três classes de genes homeóticos interagem para determinar a formação de quatro verticilos que constituem uma flor completa.

Checagem dos conceitos 5

Quais tipos de estruturas florais você espera ver nos verticilos 1 a 4 de uma planta mutante que não produz os produtos dos genes das classes A e B?

a. Carpelos, estames, estames, carpelos.
b. Sépalas, sépalas, carpelos, carpelos.
c. Sépalas, sépalas, sépalas, sépalas
d. Carpelos, carpelos, carpelos, carpelos.

Conceitos conectantes

Comparação do desenvolvimento de *Drosophila* e de flores

Agora vamos analisar dois sistemas modelo muito diferentes de desenvolvimento: a constituição da forma e padrão do corpo nas moscas-da-fruta e o desenvolvimento das estruturas florais nos angiospermas. Apesar de suas diferenças, esses dois sistemas exibem semelhanças em como os genes controlam o desenvolvimento.

Primeiro, a formação padrão em *Drosophila* e o desenvolvimento da flor nas plantas são controlados por numerosos genes que interagem de formas complexas. Por exemplo, observamos em *Drosophila* como um grande complexo de genes define sucessivamente regiões cada vez menores do embrião da mosca-da-fruta e como os genes em um nível estimulam e inibem os genes em outros níveis. Da mesma forma, o desenvolvimento floral é controlado pelos genes da classe A, classe B e classe C. A ação de cada classe depende de quais produtos de outras classes estão presentes, e os genes individuais de cada classe interagem de formas complexas para controlar a diferenciação de cada verticilo de uma flor.

Outra característica comum do desenvolvimento nas moscas-da-fruta e das flores é que muitos dos genes envolvidos nesses processos atuam ao influenciar a expressão de outros genes. Existe uma cascata de desenvolvimento em ambos os modelos, nos quais os produtos dos genes estimulam outros genes, que por sua vez, estimulam outros genes. Muitos dos produtos dos genes são proteínas reguladoras que se ligam ao DNA e afetam a transcrição de outros genes. Por exemplo, os genes *Hox* na *Drosophila* e os genes *MADS-box* nas flores codificam fatores de transcrição importantes no desenvolvimento.

Uma última semelhança é que cada sistema tem genes homeóticos, que definem a identidade de estruturas ou segmentos específicos. A mutação desses genes homeóticos produz estruturas que estão no lugar errado, como as pernas onde as antenas são normalmente encontradas nas moscas-da-fruta ou carpelos onde as sépalas ocorrem nas flores.

22.4 A morte celular programada é uma parte integral do desenvolvimento

Uma questão importante sobre o desenvolvimento é a morte das células. A morte celular define muitas partes do corpo no curso do desenvolvimento: é responsável pelo desaparecimento da cauda do girino durante a metamorfose e remove o tecido entre os dígitos para produzir a mão humana. Ela também é usada para eliminar células perigosas que escaparam dos controles naturais (ver seção sobre mutações no controle do ciclo celular e câncer no *Capítulo 23*).

Apoptose. A morte celular nos animais é iniciada pela própria célula na **apoptose**, ou suicídio celular. Nesse processo, o DNA de uma célula é degradado, seu núcleo e seu citoplasma encolhem e a célula é fagocitada por outras células sem extravasamento de seu conteúdo (**Figura 22.14 A**). As células que estão lesionadas, por outro lado, morrem de forma relativamente não controlada, chamada de necrose. Nesse processo, uma célula aumenta de volume e “explode”, derramando seu conteúdo nas células vizinhas e iniciando uma resposta inflamatória (**Figura 22.14 B**). A apoptose é essencial para a embriogênese; a maioria dos animais multicelulares não termina o desenvolvimento se esse processo for inibido.

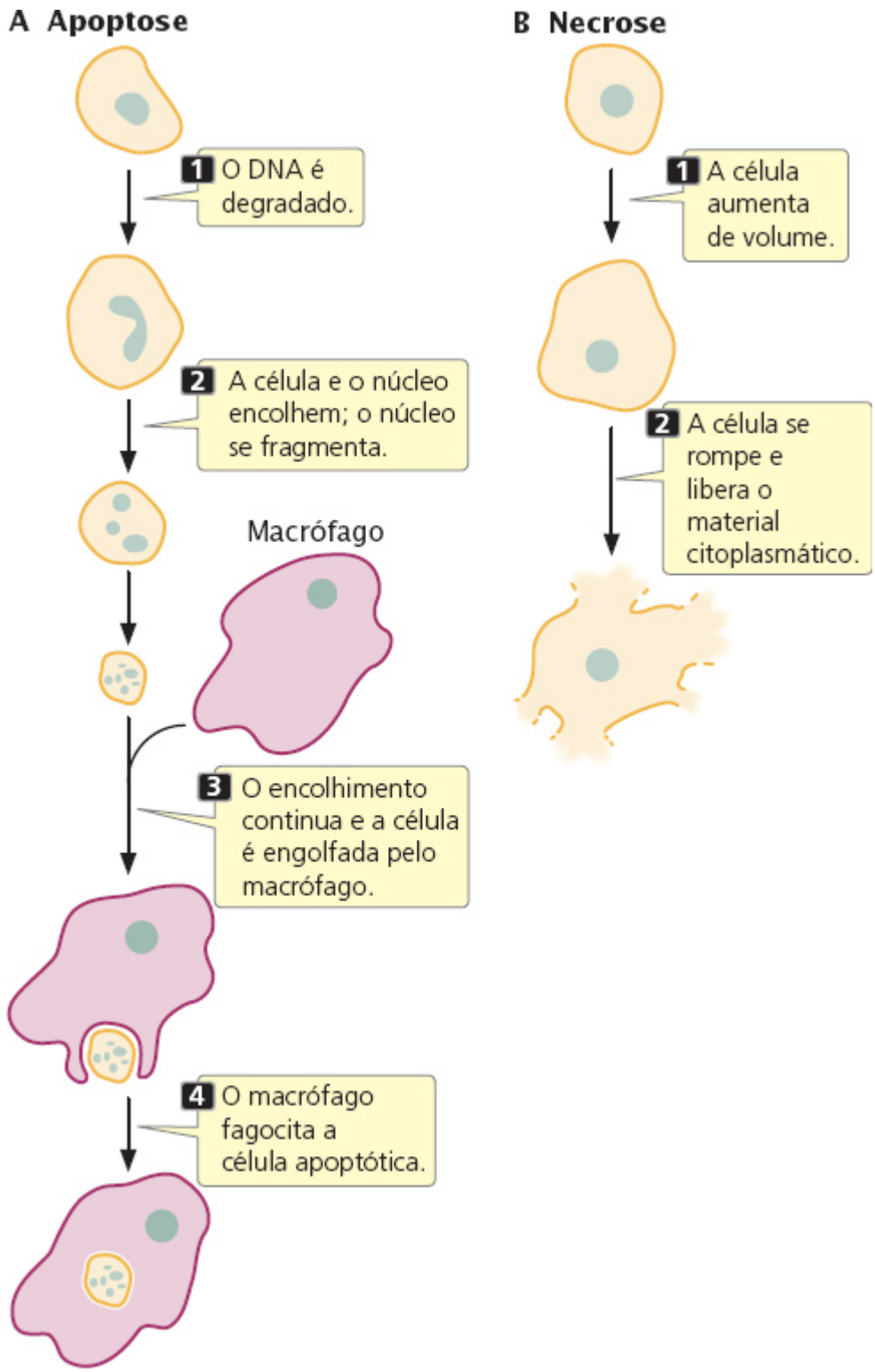


Figura 22.14 A morte celular programada pela apoptose é diferente da morte celular não controlada pela necrose.

Regulação da apoptose. De forma surpreendente, a maioria das células é programada para sofrer apoptose e sobreviverão apenas se o programa de morte interna for mantido em suspenso de forma contínua. A apoptose é altamente regulada e depende de numerosos sinais de dentro e de fora da célula. Os geneticistas identificaram muitos genes que participam de vários estágios da regulação da apoptose. Alguns desses genes codificam enzimas chamadas **caspases**, que clivam outras proteínas em sítios específicos (após o ácido aspártico). Cada caspase é sintetizada como um grande precursor inativo (uma pró-caspase) que é ativado por clivagem, frequentemente por outra caspase. Quando uma caspase é ativada, ela cliva outras pró-caspases que acionam mais atividade da caspase. A cascata resultante de atividade da caspase consequentemente cliva as proteínas essenciais para a função da célula, como as que sustentam a membrana nuclear e o citoesqueleto. As caspases também clivam uma proteína que normalmente mantém uma enzima que degrada o DNA (DNase) em uma forma inativa. A clivagem dessa proteína ativa a DNase e leva a clivagem do DNA celular, e consequente morte celular.

As pró-caspases e outras proteínas necessárias para a morte celular são produzidas de forma contínua pelas células saudáveis; dessa forma, o potencial para suicídio celular está sempre presente. Vários sinais diferentes podem disparar a apoptose; por

exemplo, a infecção por um vírus pode ativar as células imunológicas a secretar substâncias em uma célula infectada, fazendo com que a célula sofra apoptose. Acredita-se que esse processo seja um mecanismo de defesa projetado para evitar a reprodução e disseminação de vírus. Da mesma forma, a lesão ao DNA pode induzir apoptose e evitar a replicação de sequências mutantes. A lesão às mitocôndrias e o acúmulo de proteínas dobradas incorretamente no retículo endoplasmático também estimulam a morte celular programada.

Apoptose no desenvolvimento. A apoptose é importante para o desenvolvimento. À medida que os animais se desenvolvem, são produzidas células em excesso que são posteriormente eliminadas por apoptose para produzir o número adequado de células necessário para um órgão ou tecido. Em alguns casos, são criadas estruturas inteiras que são removidas posteriormente pela apoptose. Por exemplo, embriões iniciais de mamíferos desenvolvem ambos os ductos masculinos e femininos, mas os ductos de Wolff se degeneram nas fêmeas e os ductos de Muller se degeneram nos machos. A apoptose também é importante no desenvolvimento da função imunológica (ver Seção 22.6), em que os linfócitos que reconhecem as próprias células do corpo sofrem apoptose para que eles não ataquem os próprios tecidos. O sistema imunológico também pode induzir a apoptose nas células infectadas com vírus.

Várias células morrem durante o desenvolvimento embrionário de *Drosophila*. Três genes na *Drosophila* ativam as caspases essenciais para a apoptose: *reaper* (*rpr*), *grim* e *head involution defective* (*hid*). Os embriões com uma deleção que remove todos os três genes não exibem apoptose e morrem durante a embriogênese com excesso de células. Vários outros genes também afetam o processo de apoptose.

A apoptose também é crítica na metamorfose, o processo pelo qual as estruturas larvais são transformadas em estruturas adultas. Por exemplo, as glândulas salivares maiores da larva da mosca-da-fruta regridem durante a metamorfose. O hormônio ecdisona estimula a metamorfose, incluindo o início da apoptose. O ecdisona induz a expressão de *rpr* e *hid* e inibe a expressão de outros genes, o que leva à apoptose das células da glândula salivar.

Apoptose na doença. Os sintomas de muitas doenças e distúrbios são causados pela apoptose ou, em alguns casos, sua ausência. Nos distúrbios neurodegenerativos como a doença de Parkinson e a doença de Alzheimer, os sintomas são causados pela perda dos neurônios pela apoptose. Nos ataques cardíacos e AVCs, algumas células morrem por necrose, mas muitas outras sofrem apoptose. O câncer é estimulado por mutações nos genes que regulam a apoptose, levando à falência desse processo que, em condições normais, eliminaria as células cancerosas (ver *Capítulo 23*).

Conceitos

As células podem sofrer apoptose (morte celular programada), um processo altamente regulado que depende de enzimas chamadas caspases. A apoptose é importante no desenvolvimento animal e está associada com várias doenças.

Checagem dos conceitos 6

Qual a diferença da morte celular por apoptose da morte celular por necrose?

22.5 O estudo do desenvolvimento revela padrões e processos de evolução

"A ontogenia recapitula a filogenia" é uma frase familiar que foi cunhada nos anos 1860 pelo zoólogo alemão Ernst Haeckel para descrever sua crença, atualmente considerada uma simplificação exagerada, de que, durante o desenvolvimento (ontogenia), os organismos repetem sua história evolutiva (filogenia). De acordo com a crença de Haeckel, um embrião humano passa pelos estágios de peixe, anfíbio, réptil e mamífero antes de desenvolver os traços humanos. Os cientistas já reconhecem que os organismos não passam pelos estágios adultos de seus ancestrais durante o desenvolvimento, mas os embriões desses organismos apresentam semelhanças com uma grande frequência.

Genes em comum nas vias do desenvolvimento. Embora a ontogenia não recapitule precisamente a filogenia, atualmente, muitos biólogos evolucionistas estão estudando o desenvolvimento para compreender melhor o processo e os padrões da evolução. Às vezes chamado de "evo-devo", o estudo da evolução pela análise do desenvolvimento está revelando que os mesmos genes modelam as vias de desenvolvimento em organismos distantes. Os biólogos acreditavam que a segmentação nos vertebrados e invertebrados fosse apenas superficialmente semelhante, mas agora sabemos que na *Drosophila* e no cordato primitivo *Branchiostoma*, o gene *engrailed* divide o embrião em segmentos específicos. Um gene chamado

distalless, que cria as pernas de uma mosca-da-fruta, é importante no desenvolvimento dos anexos branquiais dos crustáceos. Esse mesmo gene também estimula o crescimento corporal de muitos outros organismos, desde os poliquetas à estrela-do-mar. Outro exemplo é o papel do *Pitx1* no controle da presença ou ausência de espinhas pélvicas no esgana-gatas, discutido na introdução deste capítulo. Esse mesmo gene é encontrado nos camundongos, em que também afeta o desenvolvimento dos membros traseiros, e nos humanos, em que as mutações desse gene estão associadas ao desenvolvimento de pé torto.

Um exemplo incrível de como os mesmos genes em organismos distantes podem modelar vias semelhantes de desenvolvimento é observado no desenvolvimento dos olhos nas moscas-da-fruta, camundongos e humanos. Walter Gehrinh e colaboradores examinaram o efeito do gene *eyeless* em *Drosophila*, necessário para o desenvolvimento adequado dos olhos da mosca-da-fruta. Eles modificaram geneticamente as células que expressam esse gene em partes da mosca onde ele não é normalmente expresso. Quando essas moscas saíram da pupa, tinham olhos nas asas, antenas e pernas. Essas estruturas não eram apenas tecidos que lembravam olhos, mas olhos completos com córnea, células de cone e fotorreceptores que respondiam à luz, embora as moscas não pudessem usar esses olhos para enxergar, porque não tinham a conexão com o sistema nervoso.

O gene *eyeless* tem correspondentes nos camundongos e humanos que afetam o desenvolvimento dos olhos de mamíferos. Existe outra semelhança incrível entre o gene *eyeless* da *Drosophila* e o gene *Small eye* presente nos camundongos. Nos camundongos, uma mutação em uma cópia do *Small eye* produz olhos pequenos; um camundongo homozigoto para a mutação de *Small eye* não tem olhos. Também existe uma semelhança entre o gene *eyeless* na *Drosophila* e o gene *Aniridia* nos humanos, uma mutação no *Aniridia* produz um olho humano com grave malformação. As semelhanças nas sequências de *eyeless*, *Small eye* e *Aniridia* sugerem que todos os três genes evoluíram a partir de uma sequência ancestral comum e uma via comum sustenta o desenvolvimento dos olhos nas moscas, camundongos e humanos. Essa possibilidade é surpreendente, porque se acreditava que os olhos dos insetos e mamíferos tivessem evoluído de forma independente.

Genes semelhantes podem ser parte de uma via do desenvolvimento comum a duas espécies diferentes, mas com efeitos bem diferentes. Por exemplo, um gene *Hox* chamado *AbdB* ajuda a definir a extremidade posterior de um embrião de *Drosophila*; um grupo semelhante de genes nas aves divide a asa em três segmentos. Em outro exemplo, o gene *sog* nas moscas-da-fruta estimula as células a assumir uma orientação ventral no embrião, mas a expressão de um gene semelhante chamado *chordin* nos vertebrados faz as células assumirem uma orientação dorsal, exatamente o oposto do caso das moscas-da-fruta. Nos vertebrados, os genes *toll* codificam proteínas chamadas receptores semelhantes a Toll que se ligam a moléculas nos patógenos e estimulam o sistema imunológico. Nas moscas-da-fruta, o gene *toll* atua da mesma forma na imunidade, mas também codifica uma proteína que ajuda a determinar o eixo dorsoventral, como já foi mencionado. O tema que surge desses estudos é que um pequeno conjunto comum de genes pode sustentar muitos processos do desenvolvimento básicos em muitos organismos diferentes.

Evolução pela mudança na expressão gênica. Outro conceito revelado pelos estudos em evo-devo é que muitas das principais adaptações evolutivas não são obtidas por mudanças nos tipos de proteínas produzidas, mas por mudanças na expressão de genes que codificam proteínas que regulam o desenvolvimento. Essa situação é ilustrada na introdução deste capítulo, em que a deleção de um acentuador que estimula o gene *Pitx1* nos esgana-gatas é responsável pela evolução da redução da espinha pélvica nas populações desse peixe na água doce.

Esse conceito também é observado na evolução do peixe-cego. O tetra mexicano, *Astyanax mexicanus*, ocorre normalmente em águas superficiais de riachos e rios no Texas e no norte do México. Há 10.000 anos, alguns desses peixes migraram para as cavernas. Na escuridão total do ambiente da caverna, a visão era dispensável e com o passar do tempo, eles perderam seus olhos. Atualmente, algumas 30 populações diferentes de tetras mexicanos são totalmente cegas ou sem olhos, enquanto as populações da mesma espécie que residem na superfície têm o desenvolvimento normal dos olhos.

Como os tetras mexicanos perderam seus olhos? Os zigotos do tetra mexicano começam a desenvolver os olhos, como seus primos de água rasa, mas cerca de 24 horas após a fertilização, o desenvolvimento de olhos é abolido e as células que eram destinadas para serem o cristalino morrem espontaneamente. A ausência de cristalino evita que outros componentes do olho, como a córnea e a íris, se desenvolvam. O cálice óptico e a retina se formam, mas seu crescimento é retardado e as células fotorreceptoras não se diferenciam. O olho degenerado afunda na órbita e é coberto por um retalho de pele.

Os tetras mexicanos cegos têm os mesmos genes que os tetras mexicanos de águas rasas; sua diferença está na expressão gênica. Dois genes chamados *sonic hedgehog* (*shh*) e *tiggy-winkle hedgehog* (*twhh*) são mais expressos no primórdio do olho do peixe-cego do que no peixe de superfície. (O gene *hedgehog* original foi nomeado para um fenótipo mutante em *Drosophila*, na qual o embrião mutante é coberto por dentículos como um ouriço. O gene *sonic hedgehog* foi usado por causa do personagem de desenho, e o gene *tiggy-winkle hedgehog* é nomeado em homenagem a Mrs. Tiggy-Winkle, uma personagem dos livros de Beatriz Potter.) A expressão expandida de *shh* e *twhh* ativa a transcrição de outros genes, o que provoca a apoptose das células do cristalino e o cristalino se degenera.

Quando os geneticistas injetaram o mRNA *twhh* ou *shh* nos embriões do peixe de água rasa, o desenvolvimento do cristalino parou, e os adultos que surgiram a partir desses embriões não tinham olhos. Quando foram usados fármacos para inibir a

expressão de *twhh* e *shh* nos embriões desse peixe, o desenvolvimento foi parcialmente restaurado. Esses resultados demonstram que o desenvolvimento dos olhos nos tetras mexicanos é regulado pela expressão precisa de *shh* e *twhh*. A expressão excessiva de um ou ambos os genes no peixe-cego induz à morte das células do cristalino e aborta o desenvolvimento normal dos olhos. Um pequeno aumento na transcrição do gene durante o desenvolvimento embrionário resulta em uma mudança anatômica maior no peixe adulto, mudança que permite que o peixe se adapte à escuridão do ambiente da caverna.

Resolva o Problema 27

Outro exemplo de diferenças na expressão gênica que produz mudança evolutiva é observado nos tentilhões de Darwin, um grupo de 14 aves próximas encontradas nas Ilhas Gálapos (ver Figura 26.8). Os pássaros diferem principalmente no tamanho e formato de seus bicos: os tentilhões de terra têm bicos profundos e amplos, os tentilhões do cacto têm bicos longos e pontudos e os tentilhões toutinegros têm bicos afiados e finos. Essas diferenças estão associadas com a dieta e as mudanças evolutivas no formato e no tamanho do bico ocorreram no passado quando as mudanças climáticas produziram alteração na abundância de tipos de alimentos.

Para investigar a base genética envolvida nessas mudanças evolutivas, Arkaht Abzhanov e colaboradores usaram microarranjos (ver *Capítulo 20*) para examinar as diferenças nos níveis de transcrição de muitos milhares de genes em cinco espécies de tentilhões de Darwin. Eles encontraram as diferenças na expressão de um gene que codifica uma proteína chamada calmodulina (CaM); o gene que codifica CaM foi mais expresso no bico longo e pontudo dos embriões do tentilhão cacto do que nos bicos das outras espécies. CaM participa de um processo chamado sinalização do cálcio, conhecido por afetar muitos aspectos do desenvolvimento. Quando Abzhanov e seus colaboradores ativaram a sinalização do cálcio nos embriões em desenvolvimento da galinha, as galinhas tinham bicos mais longos, como os bicos do tentilhão do cacto. Assim, esses pesquisadores puderam reproduzir, em parte, a diferença evolutiva que diferencia os tentilhões do cacto. Esse experimento mostra que as mudanças na expressão de um único gene no curso do desenvolvimento podem produzir diferenças anatômicas significativas nos adultos.

Nesses estudos e em outros, os esforços combinados dos biólogos de desenvolvimento, geneticistas e biólogos evolutivos são fontes de visões importantes de como a evolução ocorre. Embora a frase eufônica de Haeckel “a ontogenia recapitula a filogenia” estivesse errada, a evo-devo está provando que o desenvolvimento pode revelar muito sobre o processo de evolução.

22.6 O desenvolvimento da imunidade pelo rearranjo genético

Como já observamos na nossa análise sobre o desenvolvimento dos animais e das plantas, um princípio básico da biologia do desenvolvimento é que cada célula somática carreia um conjunto idêntico de informação genética e nenhum gene é perdido no desenvolvimento. Embora esse princípio se aplique à maioria das células, existem exceções importantes, uma das quais envolve os genes que codificam a função imunológica nos vertebrados. No desenvolvimento da imunidade, segmentos individuais de alguns genes são rearranjados em diferentes combinações, produzindo células imunológicas com informação genética diferente e cada uma adaptada a atacar uma substância estranha específica. Esse rearranjo e a perda de material genético são fundamentais para o poder de nossos sistemas imunológicos em nos proteger contra quase todas as substâncias estranhas concebíveis.

O sistema imunológico fornece proteção contra infecção por bactérias, vírus, fungos e parasitas. O foco de uma resposta imunológica é o **antígeno**, definido como qualquer molécula (em geral uma proteína) que provoca uma reação imunológica. O sistema imunológico é extraordinário na sua capacidade em reconhecer uma quantidade quase ilimitada de potenciais antígenos. O corpo é cheio de proteínas, então é essencial que o sistema imunológico seja capaz de distinguir entre os autoantígenos e os antígenos estranhos. Em alguns casos, a capacidade de fazer essa distinção está comprometida, e o corpo produz uma reação imunológica contra seus próprios antígenos, resultando em uma **doença autoimune** (**Quadro 22.5**).

Quadro 22.5 Exemplos de doenças autoimunes.

Doença	Tecidos atingidos
Doença de Graves, tireoidite de Hashimoto	Glândula tireoide
Febre reumática	Músculo cardíaco
Lúpus eritematoso sistêmico	Articulações, pele e outros órgãos
Artrite reumatoide	Articulações
Diabetes melito dependente de insulina	Células produtoras de insulina no pâncreas

Esclerose múltipla	A bainha de mielina que circunda os neurônios

Organização do sistema imunológico

O sistema imunológico tem muitos componentes diferentes e usa vários mecanismos para fornecer proteção contra patógenos, mas a maioria das respostas imunológicas pode ser agrupada em duas classes principais: imunidade humoral e imunidade celular. Embora seja conveniente pensar nessas classes como sistemas separados, elas interagem e influenciam uma a outra de forma expressiva.

Imunidade humoral. A função imunológica é feita por células do sangue especializadas chamadas linfócitos, que são um tipo de célula branca do sangue. A **imunidade humoral** se concentra na produção de anticorpos por linfócitos especializados chamados de **linfócitos B** (**Figura 22.15**), que amadurecem na medula óssea. Os **anticorpos** são proteínas que circulam no sangue e em outros fluidos corporais, se ligando a antígenos específicos e os marcando para serem destruídos pelas células fagocitárias. Os anticorpos também ativam um conjunto de proteínas chamadas de complemento que ajuda a romper as células e atrair os macrófagos.

Imunidade celular. Os **linfócitos T** (ver Figura 22.15) são linfócitos especializados que amadurecem no timo e respondem apenas a antígenos encontrados nas superfícies das próprias células do corpo. Esses linfócitos são responsáveis pelo segundo tipo de resposta imunológica, a **imunidade celular**.

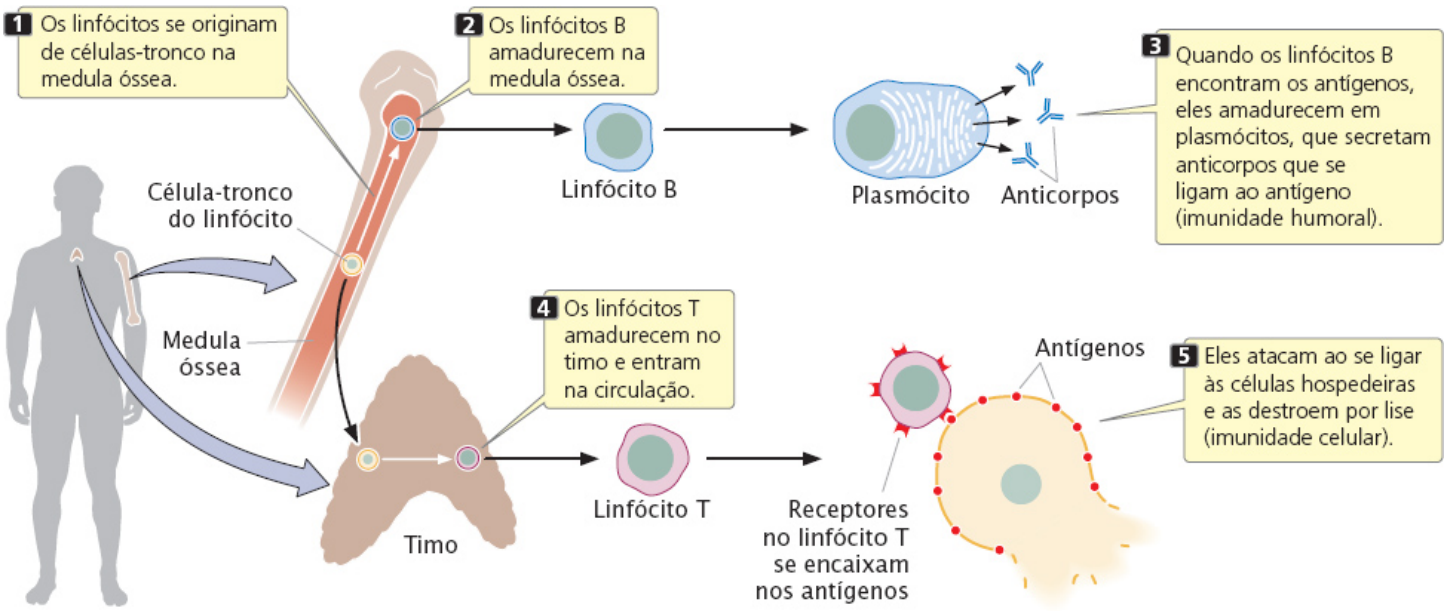


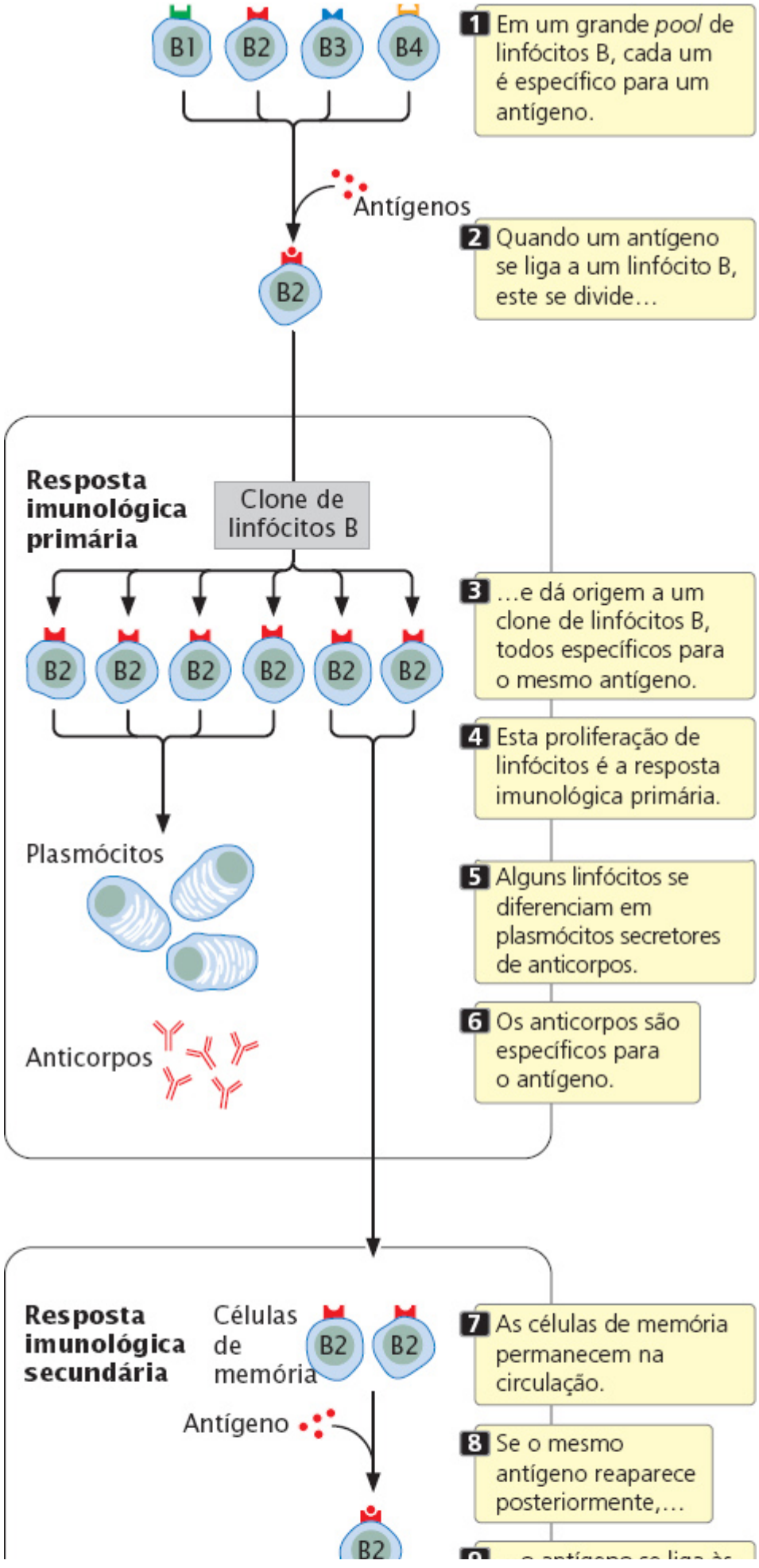
Figura 22.15 As respostas imunológicas são classificadas como imunidade humoral, na qual os anticorpos são produzidos pelos linfócitos B e a imunidade celular, que é produzida pelos linfócitos T.

Após um patógeno como um vírus ter infectado uma célula hospedeira, alguns antígenos virais surgem na superfície da célula. As proteínas, chamadas de **receptores de linfócito T**, na superfície dos linfócitos T se ligam a esses antígenos e marcam a célula infectada para destruição. Os receptores de linfócito T precisam se ligar simultaneamente a um antígeno estranho e a um autoantígeno chamado **antígeno do complexo principal de histocompatibilidade (MHC)** existente na superfície da célula (discutido posteriormente neste capítulo). Nem todos os linfócitos T atacam células com antígenos estranhos; alguns ajudam a regular as respostas imunológicas, fornecendo comunicação entre os diferentes componentes do sistema imunológico.

Seleção clonal. Como o sistema imunológico consegue reconhecer um número quase ilimitado de antígenos estranhos? De forma incrível, cada linfócito maduro é geneticamente programado para atacar um e apenas um antígeno específico: cada linfócito B maduro produz anticorpos contra um único antígeno e cada linfócito T é capaz de atacar apenas um tipo de antígeno estranho.

Se cada linfócito é específico para apenas um tipo de antígeno, como se desenvolve uma resposta imunológica? A **teoria da seleção clonal** afirma que, inicialmente, existe um grande *pool* de milhões de diferentes linfócitos, cada um capaz de se ligar a apenas um antígeno (**Figura 22.16**), então milhões de diferentes antígenos podem ser detectados. Para ilustrar a seleção clonal, vamos imaginar que uma proteína estranha penetra no corpo. Apenas alguns linfócitos no *pool* serão específicos para esse antígeno estranho. Quando um desses linfócitos encontra o antígeno estranho e se liga a ele, o linfócito é estimulado a se

dividir. O linfócito prolifera rapidamente, produzindo uma grande população de células geneticamente idênticas, um clone, cada qual específico para esse antígeno.



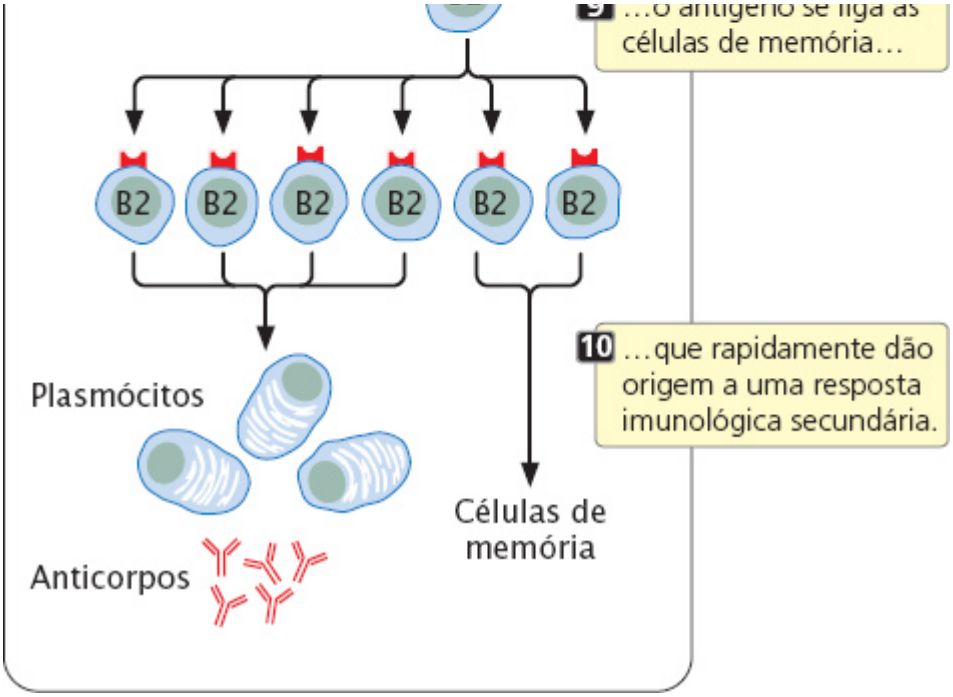


Figura 22.16 **Uma resposta imunológica a um antígeno específico é produzida pela seleção clonal.**

Essa proliferação inicial de linfócitos B e T específicos para antígeno é conhecida como **resposta imunológica primária** (ver Figura 22.16); na maioria dos casos, a resposta primária destrói o antígeno estranho. Após a resposta imunológica primária, a maioria dos linfócitos no clone morre, mas alguns continuam a circular no corpo. Essas **células de memória** podem permanecer na circulação por anos ou até pelo resto da vida da pessoa. Se o mesmo antígeno reaparecer em algum momento no futuro, as células de memória específicas para esse antígeno são ativadas e rapidamente dão origem a outro clone de células capazes de se ligar ao antígeno. O surgimento desse segundo clone é chamado de **resposta imunológica secundária** (ver Figura 22.16). A capacidade de produzir rapidamente um segundo clone de células específicas para o antígeno permite a imunidade de longa duração que se segue à recuperação de uma doença. Por exemplo, pessoas que tiveram catapora geralmente desenvolvem imunidade de longa duração para a doença. A resposta imunológica secundária também é a base para a vacinação, que estimula uma resposta imunológica primária a um antígeno e faz com que as células da memória possam produzir uma resposta secundária rapidamente se aquele mesmo antígeno surgir no futuro. Três conjuntos de proteínas são usados nas respostas imunológicas: anticorpos, receptores de linfócito T e os antígenos da histocompatibilidade principal. A seção a seguir explora como é gerada a imensa diversidade nessas proteínas.

Conceitos

Cada linfócito B e linfócito T do sistema imunológico é geneticamente capaz de se ligar a um tipo de antígeno estranho. Quando um linfócito se liga a um antígeno, o linfócito sofre repetidas divisões, dando origem a um clone de linfócitos geneticamente idênticos (a resposta primária), todos os quais são específicos para aquele mesmo antígeno. As células de memória permanecem em circulação por longos períodos de tempo; se o antígeno reaparece, as células de memória proliferam rapidamente e geram uma resposta imunológica secundária.

Estrutura da imunoglobulina

Os produtos principais da resposta imunológica humoral são os anticorpos, também chamados de imunoglobulinas. Cada molécula de imunoglobulina (Ig) é composta por quatro cadeias de polipeptídios, duas cadeias leves idênticas e duas cadeias pesadas idênticas, que formam uma estrutura com formato de Y (**Figura 22.17**). As pontes dissulfeto ligam as duas cadeias pesadas na haste do Y e prendem uma cadeia leve a uma cadeia pesada em cada braço do Y. Os sítios de ligação para antígenos estão nas extremidades dos dois braços.

Existem dois tipos básicos de cadeias leves de imunoglobulina: cadeias kappa e cadeias lambda. Uma molécula de imunoglobulina pode ter duas cadeias kappa ou duas cadeias lambda, mas ela não pode uma de cada tipo. As cadeias leves e pesadas têm uma região variável em uma extremidade e uma região constante na outra; as regiões variáveis de diferentes moléculas de imunoglobulina variam na sequência de aminoácidos, enquanto as regiões constantes de diferentes

imunoglobulinas têm sequências semelhantes. As regiões variáveis das cadeias leve e pesadas compõem as regiões de ligação ao antígeno e especificam o tipo de antígeno ao qual o anticorpo pode se ligar.

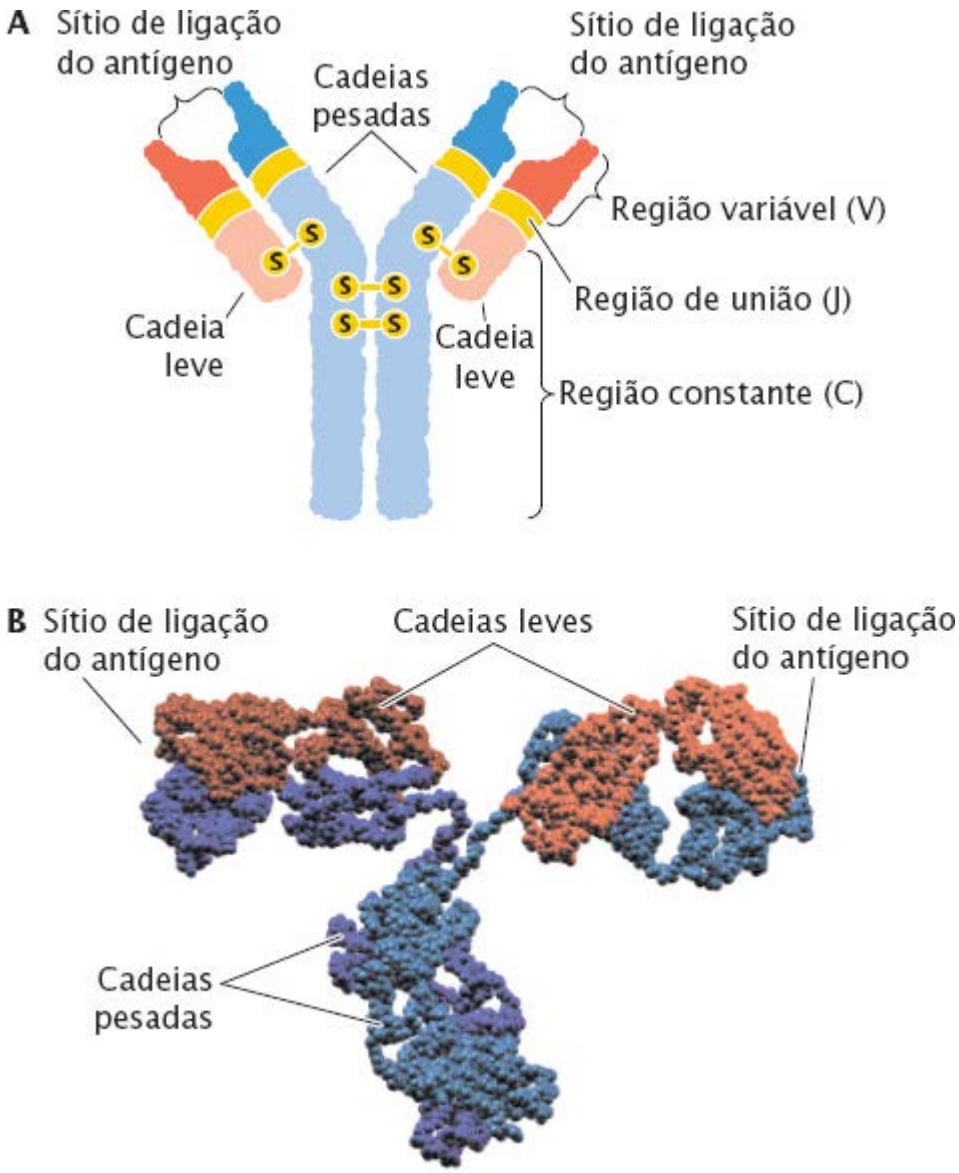


Figura 22.17 Cada molécula de imunoglobulina é composta por quatro cadeias de polipeptídios, duas cadeias leves e duas cadeias pesadas, que se combinam para formar uma estrutura em Y. A. Estrutura de uma imunoglobulina. **B.** Modelo dobrado, de espaço preenchido.

Geração da diversidade dos anticorpos

O sistema imunológico é capaz de criar anticorpos contra praticamente qualquer antígeno que possa ser encontrado na vida de uma pessoa: cada pessoa é capaz de criar cerca de 10^{15} diferentes moléculas de anticorpos. Os anticorpos são proteínas, então as sequências de aminoácidos de todos os 10^{15} potenciais anticorpos devem ser codificadas no genoma humano. Entretanto, existem menos de 1×10^5 genes no genoma humano e na verdade, apenas 3×10^9 pares de bases no total, então como essa imensa diversidade de anticorpos pode ser codificada?

A resposta está no fato de que os genes dos anticorpos são compostos por segmentos. Existem várias cópias de cada tipo de segmento, cada um discretamente diferente dos outros. Na maturação de um linfócito, os segmentos são unidos para criar um gene da imunoglobulina. A cópia de cada segmento usado é aleatória e, como existem múltiplas cópias de cada tipo, há muitas combinações possíveis dos segmentos. Um número limitado de segmentos pode codificar uma grande diversidade de anticorpos.

Para ilustrar esse processo de montagem de anticorpo, vamos analisar as cadeias leves da imunoglobulina. As cadeias kappa e lambda são codificadas por genes separados em diferentes cromossomos. Cada gene é composto por três tipos de segmentos: *V*, para variável, *J* para união e *C* para constante. Os segmentos *V* codificam a maior parte da região variável das cadeias leves, o segmento *C* codifica a região constante da cadeia, e os segmentos *J* codificam um conjunto curto de nucleotídios que unem os segmentos *V* e *C*. O número de segmentos *V, J* e *C* difere entre as espécies. Para o gene do kappa humano, existem de 30 a 35

segmentos diferentes de gene *V* funcionais, 5 diferentes genes *J* e um único segmento do gene *C*, todos presentes no DNA de linhagem celular (**Figura 22.18 A**). Os segmentos de gene *V*, que têm 400 pb de comprimento, estão localizados no mesmo cromossomo e separados um do outro em cerca de 7.000 pb. Os segmentos do gene *J* têm 30 pb de comprimento e todos juntos incluem cerca de 1.400 pb.

Inicialmente, um linfócito imaturo herda todos os segmentos de gene de *V* e todos os segmentos de gene *J* presentes na linhagem celular. Na maturação do linfócito, a **recombinação somática** em um único cromossomo desloca um dos segmentos do gene *V* para uma posição próxima a um dos segmentos de gene *J* (**Figura 22.18 B**). Na Figura 22.18 B, V_2 (o segundo de aproximadamente 35 segmentos de gene *V*) sofre recombinação somática, que ocorre próximo de J_3 (o terceiro dos 5 segmentos do gene *J*); os segmentos intervenientes são perdidos.

Após ocorrer a recombinação somática, o gene da cadeia leve é transcrito em pré-mRNA que tem um gene *V* e vários genes *J*, junto com o gene *C* (**Figura 22.18 C**). O pré-mRNA resultante é processado (**Figura 22.18 D**) para produzir um mRNA maduro com apenas sequências de um único segmento de *V, J* e *C*; esse mRNA é transcrito em uma cadeia leve funcional (**Figura 22.18 E**). Dessa forma, cada linfócito B humano maduro produz um tipo único de cadeia leve kappa e diferentes linfócitos B produzem cadeias kappa discretamente diferentes, dependendo da combinação dos segmentos *V* e *J* que são unidos.

O gene que codifica a cadeia leve lambda é organizado de forma semelhante, mas é diferente do gene kappa no número de cópias dos diferentes segmentos. A recombinação somática ocorre entre os segmentos na mesma forma que no gene kappa, gerando muitas combinações possíveis de cadeias leves lambda. O gene que codifica a cadeia pesada da imunoglobulina é organizado nos segmentos *V, J* e *C*, mas esse gene também tem segmentos *D* (para diversidade). Assim, são possíveis muitos tipos diferentes de cadeias leves e pesadas.

A recombinação somática é produzida pelas proteínas RAG1 e RAG2, que geram quebras de fita dupla em sequências de nucleotídios específicas chamadas de sequências de sinalização de recombinação que flanqueiam os segmentos dos genes *V, D, J* e *C*. As proteínas de reparo do DNA então processam e unem as extremidades dos segmentos específicos.

Além da recombinação somática, outros mecanismos auxiliam na diversidade do anticorpo. Primeiro, cada tipo de cadeia leve pode potencialmente se combinar com cada tipo de cadeia pesada para criar uma molécula funcional de imunoglobulina, aumentando o número de possíveis variações nos anticorpos. Segundo, o processo de recombinação que une os segmentos de genes *V, J, D* e *C* no linfócito B em desenvolvimento é impreciso, e alguns nucleotídios aleatórios são perdidos ou ganhos com frequência nas junções dos segmentos recombinados. Essa **diversidade juncional** acentua muito a variação entre os anticorpos. Um terceiro mecanismo que aumenta a diversidade dos anticorpos é a **hipermutação somática**, um processo que leva a alta taxa de mutação nos genes dos anticorpos. Esse processo é iniciado quando as bases citosina são desaminadas, convertendo-as em uracila. As bases uracila são eliminadas e substituídas pelos mecanismos de reparo do DNA (ver *Capítulo 18*) que são suscetíveis a erros e substituem a citosina original por uma base diferente, levando à mutação.

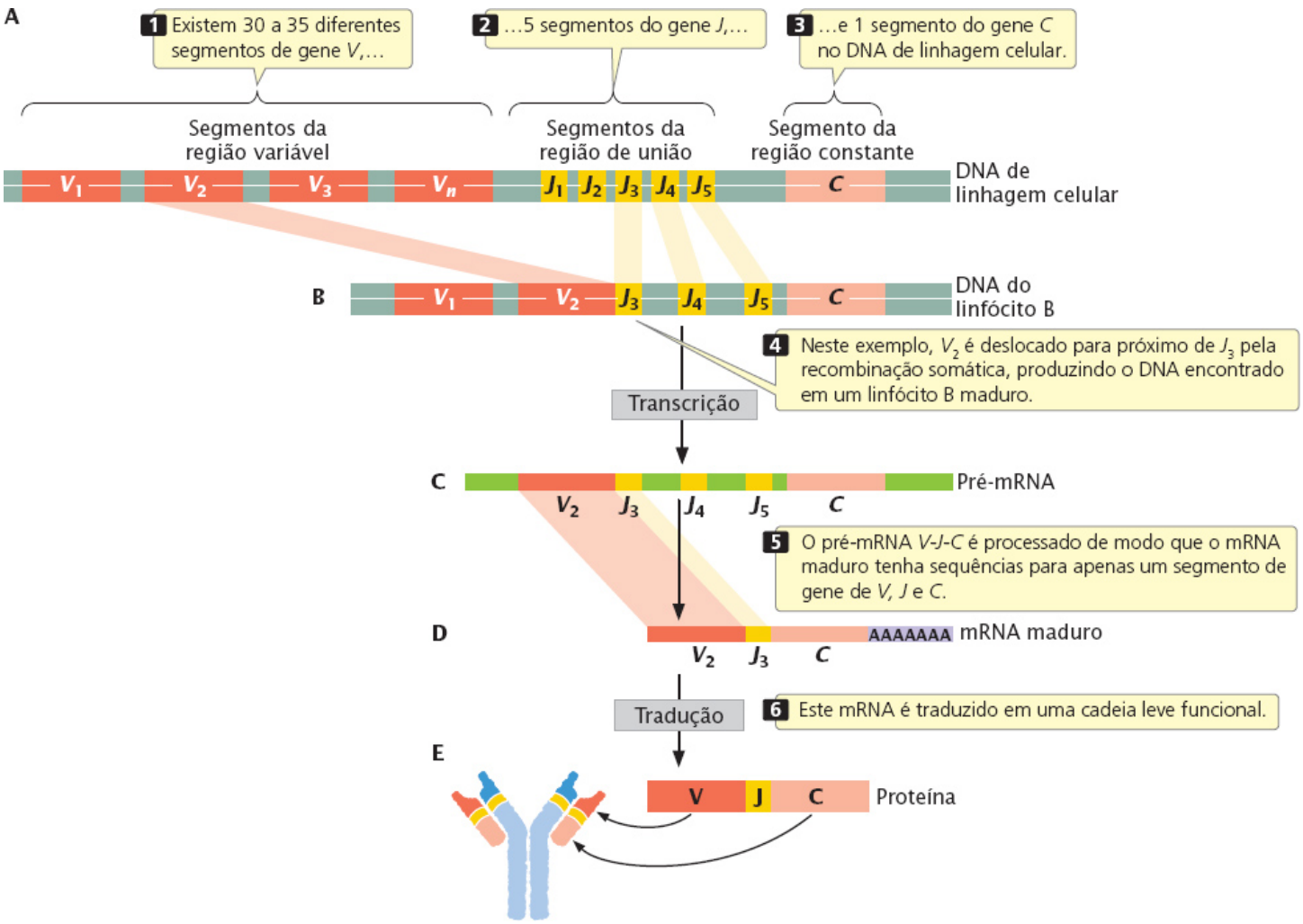


Figura 22.18 A diversidade dos anticorpos é produzida pela recombinação somática. O esquema representado aqui é a recombinação entre os segmentos de genes que codificam a cadeia kappa leve humana.

Pelos processos de recombinação somática, diversidade juncional e hipermutação somática, cada linfócito ganha um conjunto único de informação genética (diferente de outra nos outros linfócitos). Resolva o Problema 28

Conceitos

Os genes que codificam as cadeias de anticorpos estão organizados em segmentos, e o DNA de linhagem celular tem múltiplas versões de cada segmento. As muitas possíveis combinações dos segmentos *V, J* e *D* permitem uma imensa variedade de diferentes anticorpos a ser gerada. Essa diversidade é aumentada por diferentes combinações das cadeias leve e pesada, a adição e a deleção aleatória de nucleotídios nas junções dos segmentos e as altas taxas de mutação nos genes da imunoglobulina.

Checagem dos conceitos 7

Qual é a diferença da recombinação somática da recomposição alternativa do RNA?

Diversidade do receptor de linfócito T

Como os linfócitos B, cada linfócito T maduro tem a especificidade geneticamente determinada por um tipo de antígeno que é mediada pelos receptores da célula. Os receptores de linfócito T são estruturalmente semelhantes às imunoglobulinas (**Figura 22.19**) e estão localizados na superfície da célula; a maioria dos receptores de linfócito T é composta por uma cadeia de polipeptídios alfa e beta unidos por pontes dissulfeto. Cada extremidade de cada cadeia está incrustada na membrana celular; a outra extremidade se projeta para fora da célula e se liga a antígenos. Como as cadeias de imunoglobulina, cada cadeia do receptor de linfócito T tem uma região constante e uma região variável (ver Figura 22.17), as regiões variáveis das duas cadeias fornecem o sítio de ligação de antígeno.

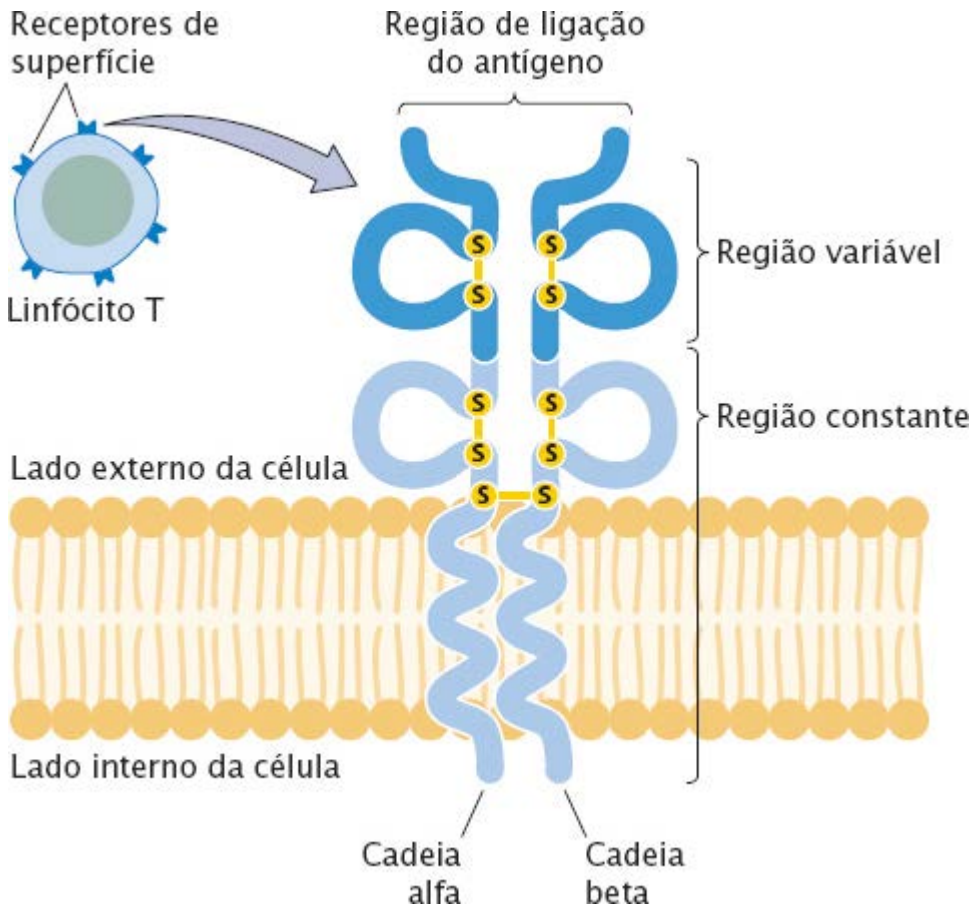


Figura 22.19 Um receptor de linfócito T é composto por duas cadeias de polipeptídios, cada uma com uma região variável e uma região constante. A maioria dos receptores de linfócito T é composta por cadeias de polipeptídios alfa e beta mantidos unidos por pontes dissulfeto. Cada extremidade de cada cadeia atravessa a membrana celular; a outra extremidade se projeta para fora da célula e se liga a antígenos.

Os genes codificadores das cadeias alfa e beta do receptor de linfócito T estão organizados de forma muito semelhante aos que codificam as cadeias pesada e leve das imunoglobulinas: cada gene é composto por segmentos que sofrem recombinação somática antes de o gene ser transcrito. Por exemplo, o gene humano para a cadeia alfa é composto inicialmente de 44 a 46 segmentos de gene *V*, 50 segmentos de gene *J* e um único segmento do gene *C*. A organização do gene para a cadeia beta é semelhante, exceto que ele também tem os segmentos do gene D. As cadeias alfa e beta se combinam de forma aleatória e existe a diversidade juncional, mas não existe indício para hipermutação somática nos genes do receptor de linfócito T.

Conceitos

Como os genes que codificam os anticorpos, os genes para as cadeias de receptor de linfócito T são compostos por segmentos que sofrem recombinação somática, gerando uma enorme diversidade de sítios de ligação de antígeno.

Genes do complexo principal de histocompatibilidade

Quando os tecidos são transferidos de uma espécie para outra ou de um membro para outro em uma espécie, os tecidos transplantados costumam ser rejeitados pelo hospedeiro. Os resultados dos estudos iniciais demonstraram que essa rejeição do enxerto ocorre por causa de uma resposta imunológica que ocorre quando os antígenos na superfície das células do tecido enxertado são detectados e atacados pelos linfócitos T no organismo hospedeiro. Os antígenos que estimulam a rejeição do enxerto são chamados de antígenos de histocompatibilidade e são codificados por um grupo de genes chamados de complexo principal de histocompatibilidade (MHC).

Os linfócitos T são ativados apenas quando o receptor de linfócito T se liga simultaneamente a um antígeno estranho e ao antígeno próprio de histocompatibilidade da célula hospedeira. O motivo para essa exigência não está claro; ela pode reservar os linfócitos T para ação contra patógenos que invadiram as células. Quando um corpo estranho, como um vírus, é fagocitado por um macrófago ou outra célula, os pedaços parcialmente digeridos do corpo estranho com antígenos são apresentados na superfície da célula (**Figura 22.20**). Uma célula infectada com um vírus também pode expressar antígenos virais na sua superfície. Por meio de seus receptores, os linfócitos T se ligam à proteína de histocompatibilidade e ao antígeno estranho e

secretam substâncias que destroem a célula que contém o antígeno, ativam outros linfócitos B e T ou fazem ambos.

Os genes MHC estão entre os genes mais variáveis conhecidos: existem mais de 100 diferentes alelos para cada *loci* de MHC. Como cada pessoa tem cinco ou mais *loci* MHC e como são possíveis muitos alelos em cada *locus*, duas pessoas (com a exceção dos gêmeos idênticos) não produzem o mesmo conjunto de antígenos de histocompatibilidade. A variação nos antígenos da histocompatibilidade permite que tenhamos uma identidade única para nossas próprias células, o que permite que nosso sistema imunológico diferencie o próprio (*self*) do não próprio (*nonself*). Essa variação também é a causa da rejeição dos transplantes de órgãos.

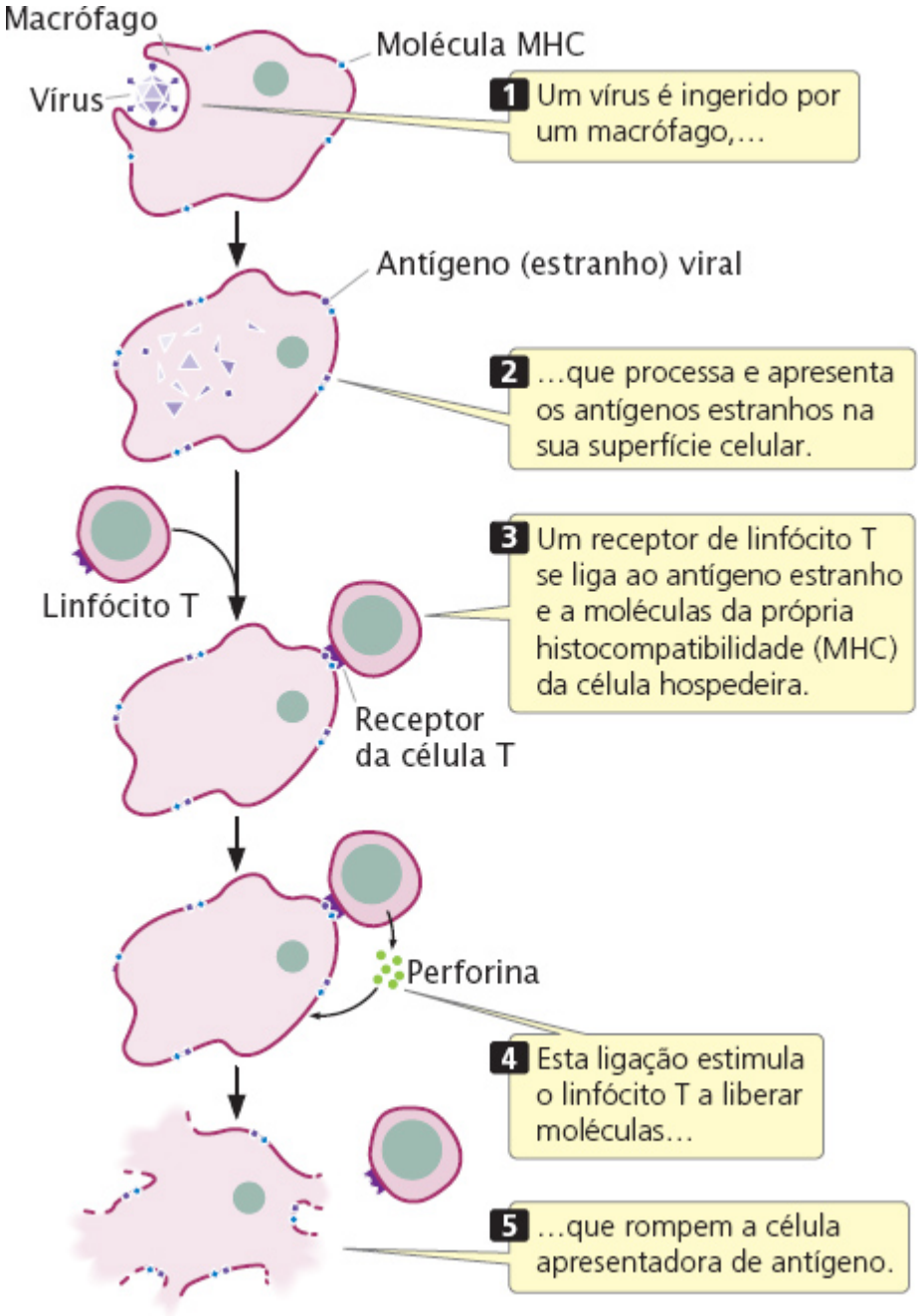


Figura 22.20 Os linfócitos T são ativados ao se ligar a um antígeno estranho e a um antígeno de histocompatibilidade na superfície de uma autocélula.

Conceitos

Os genes MHC codificam proteínas que fornecem identidade às células de cada organismo individual. Para produzir uma resposta imunológica, um receptor de linfócito T deve se ligar simultaneamente ao antígeno de histocompatibilidade (próprio) e a um antígeno estranho específico.

Genes e transplantes de órgãos

Para uma pessoa com um órgão seriamente comprometido, o transplante pode ser a única esperança de sobrevida. O transplante

bem-sucedido requer mais do que as habilidades de um cirurgião; também requer uma combinação genética entre o paciente e o doador do órgão. O destino do tecido transplantado depende muito do tipo de antígenos existentes na superfície de suas células. Como tecidos estranhos são rejeitados pelo hospedeiro, é muito difícil o transplante bem-sucedido de tecidos entre pessoas diferentes. A rejeição de tecido pode ser parcialmente inibida por fármacos que interferem com a imunidade celular. Infelizmente, esse tratamento pode criar vários problemas para os pacientes transplantados porque eles podem ter dificuldade para combater patógenos comuns e morrer de infecção. A única opção para controlar a reação imunológica é combinar cuidadosamente o doador e o receptor, maximizando as semelhanças genéticas.

Os antígenos do tecido que deflagram a reação imunológica mais forte são os mesmos usados pelo sistema imunológico para marcar suas próprias células: os codificados pelo complexo principal de histocompatibilidade. O MHC abrange uma região de mais de 3 milhões de pares de bases no cromossomo 6 humano e tem muitos alelos, fornecendo diferentes antígenos MHC nas células de pessoas diferentes e permitindo que o sistema imunológico reconheça as células estranhas. A intensidade de uma rejeição imunológica de um órgão transplantado depende do número de antígenos MHC não compatíveis nas células do tecido transplantado. Os antígenos eritrocitários do sistema AB0 também são importantes porque deflagram uma forte reação imunológica. O doador ideal é o próprio irmão gêmeo idêntico do paciente, que teria os mesmos antígenos de MHC e AB0. Infelizmente, a maioria dos pacientes não tem um gêmeo idêntico. O segundo melhor doador é um irmão com os mesmos antígenos MHC e AB0. Se não houver um irmão, são considerados os doadores da população geral. É feita uma tentativa para combinar o máximo possível de antígenos MHC do doador e do receptor, e são usados fármacos imunossupressores para controlar a rejeição por causa de incompatibilidade. O sucesso duradouro dos transplantes depende do grau de compatibilidade. As taxas de sobrevida após transplantes de rim (o maior sucesso dos transplantes de órgãos importantes) aumentam de 63% com zero ou uma combinação de MHC para 90% com quatro pareamentos.

Agora os cientistas tiveram sucesso ao induzir as células adultas a perder suas características especializadas e retornar para um estado desdiferenciado chamado de célula-tronco pluripotente (ver *Capítulo 21*), capaz de se desenvolver em muitos tipos celulares diferentes. No futuro, pode ser possível criar células-tronco pluripotentes a partir das células adultas de uma pessoa, e então cultivar essas células em tecidos ou órgãos que poderão ser transplantados de volta para a mesma pessoa. Tais células teriam os mesmos antígenos MHC que a célula original, evitando a rejeição imunológica que ocorre com os transplantes entre diferentes pessoas.

Resumo dos conceitos

- Cada organismo multicelular começa como uma única célula que tem o potencial para se desenvolver em um tipo celular. À medida que o desenvolvimento continua, a célula fica comprometida a destinos específicos. Os resultados dos experimentos de clonagem demonstraram que esse processo surge a partir de expressão gênica diferenciada
- No embrião inicial de *Drosophila*, a determinação é feita por uma cascata de controle gênico
- Os eixos dorsoventral e anteroposterior desse embrião são estabelecidos pelos genes da polaridade do ovo, que são expressos na genitora e produzem RNA e proteínas que são depositadas no citoplasma do ovo. As diferenças iniciais na distribuição dessas moléculas regulam a expressão gênica em várias partes do embrião. O eixo dorsoventral é definido por um gradiente de concentração da proteína Dorsal e o eixo anteroposterior é definido pelos gradientes de concentração das proteínas Bicoid e Nanos
- Três tipos de genes de segmentação atuam em sequência para determinar o número e a organização dos segmentos embrionários em *Drosophila*. Os genes *gap* estabelecem grandes seções do embrião, os genes da regra dos pares afetam os segmentos alternados e os genes de polaridade segmentar afetam a organização dos segmentos individuais. Os genes homeóticos então definem a identidade dos segmentos individuais da *Drosophila*
- Os genes homeóticos controlam o desenvolvimento das estruturas das flores. Três conjuntos de genes interagem para determinar a formação de quatro verticilos que constituem uma flor completa
- A apoptose, ou a morte celular programada, é um processo altamente regulado, dependente das caspases, proteínas que rompem proteínas. A apoptose é importante no desenvolvimento de muitos animais
- O sistema imunológico é a rede de defesa primária nos vertebrados. Na imunidade humoral, os linfócitos B produzem anticorpos que se ligam a antígenos estranhos, enquanto na imunidade celular, os linfócitos T atacam as células carreando antígenos estranhos
- Cada linfócito B e T é capaz de se ligar a apenas um tipo de antígeno estranho. Quando um linfócito se liga a um antígeno, o linfócito se divide e dá origem a um clone de células, cada um específico para o mesmo antígeno – a resposta imunológica primária. Algumas células de memória permanecem em circulação por longos períodos de tempo e, com a exposição ao mesmo antígeno, podem proliferar rapidamente e gerar uma resposta imunológica secundária
- As imunoglobulinas (anticorpos) são codificadas por genes compostos por vários tipos de segmentos de genes; o DNA da linhagem germinativa tem múltiplas cópias desses segmentos de genes, que apresentam discreta diferença na sequência. A

recombinação somática produz aleatoriamente uma versão de cada segmento para produzir um único gene completo, permitindo muitas combinações. A diversidade é aumentada pela adição e pela deleção aleatórias de nucleotídios nas junções dos segmentos e por uma alta taxa de mutação

- Os genes da linhagem celular para os receptores de linfócito T são compostos por segmentos com múltiplas cópias variadas. A recombinação somática gera muitos tipos diferentes de receptores de linfócito T em diferentes células. A diversidade juncional também adiciona a variabilidade do receptor de linfócito T
- O complexo principal de histocompatibilidade codifica vários antígenos de histocompatibilidade. O antígeno do MHC permite que o sistema imunológico diferencie o próprio (*self*) do não próprio (*nonself*). Cada *locus* para o MHC tem vários alelos.

Termos importantes

Anticorpo
Antígeno
Antígeno do complexo principal de histocompatibilidade (MHC)
Apoptose
Caspase
Célula de memória
Complexo *Antennapedia*
Complexo bitórax
Complexo homeótico (HOM-C)
Determinação
Diversidade juncional
Doença autoimune
Genes da polaridade do ovo
Genes da regra dos pares
Genes de polaridade segmentar
Genes de segmentação
Genes *gap*
Genes homeótico
Genes *Hox*
Hipermutação somática
Homeobox
Imunidade celular
Imunidade humoral
Linfócitos B
Linfócitos T
Morfógeno
Receptor de linfócito T
Recombinação somática
Resposta imunológica primária
Resposta imunológica secundária
Teoria da seleção clonal
Totipotência

Respostas da Checagem dos conceitos

1. Não, ele *não* prova que o material genético não é perdido durante o desenvolvimento, porque a diferenciação ainda não ocorreu em um embrião inicial. O embrião inicial provavelmente ainda teria todos seus genes e poderia dar origem a um animal completo. O uso de células especializadas, como uma célula de uma teta, prova que os genes não são perdidos durante o desenvolvimento porque se fossem perdidos não existiriam animais clonados.
2. c.
3. b.
4. d.
5. d.
6. Na morte celular por necrose, a célula aumenta de volume e estoura, provocando uma resposta inflamatória. Na morte

celular por apoptose, o DNA da célula é degradado, seu núcleo e citoplasma encolhem e a célula é fagocitada, sem extravasamento do conteúdo celular.

7. A recombinação somática ocorre pelo rearranjo de segmentos de DNA então cada linfócito tem uma sequência diferente de nucleotídios no seu DNA. A recomposição (*splicing*) alternativa (ver *Capítulo 14*) ocorre pelo rearranjo de sequências de RNA no pré-mRNA; não ocorre mudança no DNA que codifica o pré-mRNA. A geração de diversidade de anticorpos requer a recombinação somática das sequências de DNA e a recomposição alternativa das sequências de pré-mRNA.

Problemas desenvolvidos

Problema 1

Se um ovo de *Drosophila* fertilizado for perfurado na extremidade anterior e uma pequena quantidade de citoplasma extravasar, qual seria o efeito mais provável no embrião da mosca?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Prováveis efeitos no desenvolvimento com a remoção do citoplasma da extremidade anterior de um ovo de mosca fertilizado.

Quais informações são fornecidas para solucionar o problema?
O citoplasma da extremidade anterior é removido.

Para a solução deste problema, revise:
Genes da polaridade do ovo, na Seção 22.2.

Etapas para a solução

Os genes da polaridade do ovo determinam os principais eixos do desenvolvimento no embrião da *Drosophila*. Um desses genes é o *bicoid*, que é transcrito na genitora. Um mRNA *bicoid* passa para o ovo e o mRNA é ancorado na extremidade anterior do ovo. Após o ovo ser posto, o mRNA *bicoid* é traduzido na proteína Bicoid, que forma um gradiente de concentração ao longo do eixo anteroposterior do embrião. A elevada concentração da proteína Bicoid na extremidade anterior induz o desenvolvimento das estruturas anteriores como a cabeça da mosca-da-fruta. Se a extremidade anterior do ovo for perfurada, o citoplasma contendo altas concentrações da proteína Bicoid será perdido, reduzindo a concentração dessa proteína na extremidade anterior. O resultado será que o embrião não desenvolve a cabeça e as estruturas torácicas na extremidade anterior.

Lembrete: O gene *bicoid* é um gene da polaridade do ovo que ajuda a determinar o eixo anteroposterior do embrião em desenvolvimento.

Problema 2

As moléculas de imunoglobulina de determinada espécie de mamífero têm cadeias leves kappa e lambda e cadeias pesadas. O gene kappa é composto por 250 segmentos *V* e 8 segmentos *J*. O gene lambda é composto por 200 segmentos *V* e 4 segmentos *J*. O gene para a cadeia pesada tem 300 segmentos *V*, 8 *J* e 4 *D*. Se considerarmos apenas a recombinação somática e as combinações aleatórias das cadeias leves e pesadas, quantos tipos de anticorpos diferentes podem ser produzidos por essa espécie?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
O número de diferentes tipos de anticorpos que podem ser produzidos se considerarmos a recombinação somática e as combinações aleatórias das cadeias leves e pesadas.

Quais informações são fornecidas para solucionar o problema?

- O gene kappa tem 250 segmentos *V* e 8 segmentos *J*
- O gene lambda tem 200 segmentos *V* e 4 segmentos *J*

- O gene para a cadeia pesada tem 300 segmentos *V*, 8 *J* e 4 *D*.

Para a solução deste problema, revise:
Geração da diversidade dos anticorpos, na Seção 22.6.

Etapas para a solução

Para a cadeia leve kappa, existem 250 × 8 = 2.000 combinações; para a cadeia leve lambda, existem 200 × 4 = 800 combinações, então é possível um total de 2.800 tipos diferentes de cadeias leves. Para as cadeias pesadas, existem 300 × 8 × 4 = 9.600 tipos possíveis. Qualquer uma das 2.800 cadeias leves pode combinar-se com qualquer uma das 9.600 cadeias pesadas, então existem 2.800 × 9.600 = 26.880.000 diferentes tipos de anticorpos possíveis a partir da recombinação somática e da combinação aleatória apenas. A diversidade juncional e a hipermutação somática aumentariam muito essa diversidade.

Dica: Para determinar o número de diferentes tipos de anticorpos, multiplique o número de possíveis cadeias leves × o número de possíveis cadeias pesadas.

Dica: O número de cada tipo de cadeia leve é composto pelo número de segmentos *V* × o número de segmentos *J*.

Questões de compreensão

Seção 22.1

1. Que experimentos sugerem que os genes não são perdidos nem permanentemente alterados no desenvolvimento?

Seção 22.2

2. Explique resumidamente como a proteína Dorsal é redistribuída na formação do embrião da *Drosophila* e como essa redistribuição ajuda a estabelecer o eixo dorsoventral no embrião inicial.
3. Descreva resumidamente como os genes *bicoid* e *nanos* ajudam a determinar o eixo anteroposterior da mosca-da-fruta.
4. Liste as três principais classes de genes de segmentação e destaque a função de cada uma.
5. Qual o papel dos genes homeóticos no desenvolvimento das moscas-da-fruta?

Seção 22.3

6. Como os genes das classes A, B e C nas plantas trabalham juntos para determinar as estruturas da flor?

Seção 22.4

7. O que é apoptose e como ela é regulada?

Seção 22.6

8. Explique como cada um dos processos a seguir contribui com a diversidade do anticorpo.
   a. Recombinação somática.
   b. Diversidade juncional.
   c. Hipermutação.
9. Qual é a função dos antígenos do MHC? Por que os genes codificadores desses antígenos são tão variáveis?

Questões e problemas aplicados

Seção 22.1

10. Se os telômeros são normalmente encurtados após cada ciclo de replicação nas células somáticas (ver *Capítulo 12*), que previsão você faria sobre o comprimento dos telômeros em Dolly, a primeira ovelha clonada?

Seção 22.2

11. Uma substância causa a degradação da proteína Cactus. Qual seria o efeito ao administrar essa substância nos embriões de

Drosophila em desenvolvimento?

12. Qual seria o efeito de eliminar o gene *toll* nos embriões de *Drosophila*?
13. Por que as mutações em *bicoid* e *nanos* exibem efeitos genéticos maternos (uma mutação na genitora produz um fenótipo que aparece nos descendentes, ver *Capítulo 5*), mas as mutações em *runt* e *gooseberry* não exibem? (Dica: Ver Quadros 22.3 e 22.4.)

*14. Dê exemplos de genes que afetam o desenvolvimento nas moscas-da-fruta ao regular a expressão gênica no nível de (a) transcrição e (b) tradução.

15. Usando a Figura 22.5, indique em que estágio os genes de segmentação, genes homeóticos e genes da polaridade do ovo afetariam o desenvolvimento.
16. Qual seria o efeito mais provável no desenvolvimento ao perfurar a extremidade posterior de um ovo de *Drosophila*, permitindo que uma pequena quantidade de citoplasma extravase e então injetando esse citoplasma na extremidade anterior de outro ovo?

ANÁLISE DE DADOS

17. Christiane Nüsslein-Volhard e colaboradores fizeram vários experimentos em uma tentativa de compreender o que determina as extremidades anterior e posterior de uma larva de *Drosophila* (revisado em C. Nüsslein-Volhard, H. G. Frohnhofer, and R. Lehmann. 1987. *Science* 238:1675-1681). Eles isolaram moscas-da-fruta com mutações no gene *bicoid* (*bcd*$^{-}$). Essas moscas produziram embriões que não tinham cabeça nem tórax. Quando eles transplantaram o citoplasma da extremidade anterior de um ovo de uma fêmea do tipo selvagem para a extremidade anterior de um ovo de uma fêmea *bicoid* mutante, o embrião desenvolveu cabeça e tórax normais. Entretanto, transplantar o citoplasma da extremidade posterior de um ovo de uma fêmea do tipo selvagem na extremidade anterior de um ovo de uma fêmea *bicoid* não teve efeito. Explique esses resultados usando o que você sabe sobre as proteínas que controlam a determinação do eixo anteroposterior.
18. Qual seria o resultado mais provável ao injetar mRNA *bicoid* na extremidade posterior de um embrião de *Drosophila* e inibir a tradução do mRNA *nanos*?
19. Qual seria o efeito mais provável ao inibir a tradução do mRNA *hunchback* em todo o embrião?

*20. Geneticistas moleculares fizeram experimentos nos quais eles alteraram o número de cópias do gene *bicoid* nas moscas, afetando a quantidade de proteína Bicoid produzida.
   a. Qual seria o efeito no desenvolvimento de um número maior de cópias do gene *bicoid*?
   b. Qual seria o efeito no desenvolvimento de um número menor de cópias do *bicoid*? Justifique sua resposta.

21. Qual seria o efeito mais provável no desenvolvimento da mosca-da-fruta de uma deleção do gene *nanos*?

*22. Dê um exemplo de um gene encontrado em cada uma das categorias de genes (polaridade do ovo, *gap*, regra dos pares e assim em diante) listada na Figura 22.11.

23. No Capítulo 1, analisamos o pré-formacionismo, a ideia inicial sobre hereditariedade que sugeria que existia um homem minúsculo (homúnculo) dentro do ovo ou do espermatozoide com todas as características de um humano adulto em miniatura. De acordo com essa ideia, o homúnculo apenas aumentaria durante o desenvolvimento. Que tipos de evidência apresentados neste capítulo comprovam que o pré-formacionismo é falso?

Seção 22.3

24. Explique como (a) ausência de expressão do gene da classe B produz as estruturas de flores observadas nos mutantes da classe B (ver Figura 22.12 C) e (b) a ausência de produto de gene da classe C produz as estruturas observadas nos mutantes da classe C (ver Figura 22.12 D).

*25. Como você esperaria que fosse uma flor em uma planta que não tem os genes das classes A e B? Em uma planta sem os genes das classes B e C?

26. Qual seria a estrutura da flor de uma planta na qual a expressão dos seguintes genes é inibida?
   a. Expressão dos genes da classe B é inibida no segundo verticilo, mas não no terceiro verticilo.
   b. Expressão dos genes da classe C é inibida no terceiro verticilo, mas não no quarto verticilo.
   c. Expressão dos genes da classe A é inibida no primeiro verticilo, mas não no segundo verticilo.
   d. Expressão dos genes da classe A é inibida no segundo verticilo, mas não no terceiro verticilo.

Seção 22.5

ANÁLISE DE DADOS

*27. Willliam Jeffrey e colaboradores cruzaram tetras mexicanos de água rasa que tinham olhos totalmente desenvolvidos com

os tetras mexicanos cegos das cavernas. Os descendentes desse cruzamento tinham olhos pequenos e uniformes em comparação com os peixes de superfície (Y. Yamamoto, D. W. Stock, and W. R. Jeffrey. 2004. *Nature* 431:844-847). Que previsão você poderia fazer sobre a expressão do *shh* nos embriões desses descendentes com relação a sua expressão nos embriões do peixe de água rasa?

Seção 22.6

*28. Em uma determinada espécie, o gene para a cadeia leve kappa tem 200 segmentos do gene *V* e 4 segmentos *J*. No gene para a cadeia leve lambda, essa espécie tem 300 segmentos *V* e 6 segmentos *J*. Se apenas a variabilidade que surge a partir da recombinação somática for levada em consideração, quantos tipos diferentes de cadeias leves são possíveis?

29. Com base nas informações fornecidas na Figura 22.16, qual seria o provável efeito de uma mutação que evitaria a formação das células de memória?

30. No livro *Chromosome 6* de Robin Cook, uma empresa de biotecnologia fez a engenharia genética em bonobos (um tipo de chimpanzé) para usá-los como futuros doadores de órgãos para clientes. Os genes dos bonobos foram alterados de forma que não ocorreria rejeição de tecido quando seus órgãos fossem transplantados em um cliente. Quais genes precisariam ser alterados para essa ideia funcionar? Explique sua resposta.

Questões desafiadoras

Seção 22.2

31. Como aprendemos neste capítulo, a proteína Nanos inibe a tradução do mRNA *hunchback*, reduzindo a concentração da proteína Hunchback na extremidade posterior de um embrião de mosca-da-fruta e estimulando a diferenciação das características posteriores. Os resultados dos experimentos demonstraram que a ação de Nanos no mRNA *hunchback* depende da existência de uma sequência de 11 bases localizada na região 3′ não traduzida (3′ UTR) do mRNA *hunchback*. Essa sequência foi chamada de elemento de resposta Nanos (NRE). Existem duas cópias de NRE em 3′ UTR do mRNA *hunchback*. Se uma cópia de NRE for adicionada à 3′ UTR de outro mRNA produzido por um gene diferente, o mRNA agora será reprimido por Nanos. A repressão é maior se vários NREs forem adicionados. Com base nessas observações, proponha um mecanismo de inibição da tradução do mRNA *hunchback* pela proteína Nanos.

32. Considerando a distribuição dos genes *Hox* entre os animais, o que você prevê sobre o número e o tipo de genes *Hox* no ancestral comum de todos os animais?

Seção 22.6

33. A ataxia-telangiectasia (ATM) é uma rara doença neurodegenerativa genética. Cerca de 20% das pessoas com ATM desenvolvem leucemia linfocítica aguda ou linfoma, cânceres das células do sistema imunológico. As células em muitos desses cânceres exibem rearranjos cromossômicos, com rupturas de cromossomos ocorrendo nos genes de anticorpos e de receptores de linfócitos T (A. L. Bredemeyer *et al*. 2006. *Nature* 442:466-470). Muitas pessoas com ATM também têm um sistema imunológico enfraquecido, tornando-as suscetíveis a infecções respiratórias. A pesquisa mostrou que o *locus* responsável pela ATM participa do reparo das quebras de fita dupla. Explique por que as pessoas com um defeito genético no reparo de quebras de fita dupla poderiam ter maior incidência de rearranjos cromossômicos nas suas células do sistema imunológico e por que seus sistemas imunológicos poderiam estar enfraquecidos.

23 A Genética do Câncer

A paladina e a disseminação do câncer

O câncer pancreático é um dos mais graves de todos os cânceres. Com cerca de 45.000 novos casos todos os anos nos EUA, ele é apenas a décima forma mais comum da doença, mas é a quarta causa principal de morte por câncer, matando mais de 38.000 pessoas por ano. A maioria das pessoas com câncer pancreático sobrevive menos de 6 meses após o diagnóstico; apenas 5% sobrevivem mais de 5 anos. A principal razão da letalidade desse câncer é sua propensão a se disseminar rapidamente para os linfonodos e outros órgãos. A maioria dos sinais/sintomas não surge até o câncer estar avançado e ter invadido outros órgãos. Então, o que torna tão provável a disseminação do câncer pancreático?

Em 2006, os pesquisadores identificaram um gene-chave que contribui para o desenvolvimento desse câncer, uma fonte importante de discernimento sobre a natureza agressiva da doença. Os geneticistas na University of Washington em Seattle descobriram uma família singular na qual nove membros em três gerações foram diagnosticados com câncer pancreático (**Figura 23.1**). Nove membros da família tiveram crescimentos pré-câncer que tinham probabilidade de desenvolver câncer pancreático. Nessa família, esse tipo de câncer era herdado como um traço autossômico dominante.

Ao usar as técnicas de mapeamento gênico, os geneticistas determinaram que o gene responsável pelo câncer pancreático na família está localizado em uma região no braço longo do cromossomo 4. Infelizmente, esta região inclui 16 milhões de pares de bases e 250 genes.

Para determinar qual dos 250 genes na região delineada poderia ser responsável pelo câncer na família, os pesquisadores projetaram um microarranjo diferenciado (ver *Capítulo 20*), com sequências da região. Eles usaram esse microarranjo para examinar a expressão gênica nos tumores pancreáticos e crescimentos pré-cancerosos nos membros da família, assim como tumores pancreáticos esporádicos em outras pessoas e no tecido pancreático normal de pessoas não afetadas. Os pesquisadores justificaram que o gene do câncer poderia ser superexpresso ou subexpresso nos tumores em relação com o tecido normal. Os dados a partir do microarranjo revelaram que o gene mais superexpresso nos tumores pancreáticos e crescimentos pré-cancerosos foi um codificador de um componente crítico do citoesqueleto, um gene chamado *palladin*. O sequenciamento demonstrou que todos os membros da família com câncer pancreático tinham uma mutação idêntica no éxon 2 do gene *palladin*.

O gene *palladin* é uma homenagem ao arquiteto renascentista Andrea Palladio por seu papel central na arquitetura da célula. A proteína paladina atua como um arcabouço para a ligação a outras proteínas do citoesqueleto necessárias para manter o formato, o movimento e a diferenciação da célula. A capacidade de uma célula do câncer se disseminar está diretamente relacionada com seu citoesqueleto; as células que se disseminam apresentam uma arquitetura deficiente do citoesqueleto, permitindo que elas se destaquem facilmente da massa do tumor primário e migrem para outros tecidos. Para determinar se as mutações no gene *palladin* afetam a mobilidade da célula, os pesquisadores criaram células geneticamente modificadas com uma cópia mutante do gene *palladin* e testaram a capacidade de migração dessas células. As células com o gene *palladin* mutante eram 33% mais eficientes para migrar do que as células com *palladin* normal, demonstrando que o gene *palladin* contribui para a capacidade das células do câncer pancreático em se disseminar.

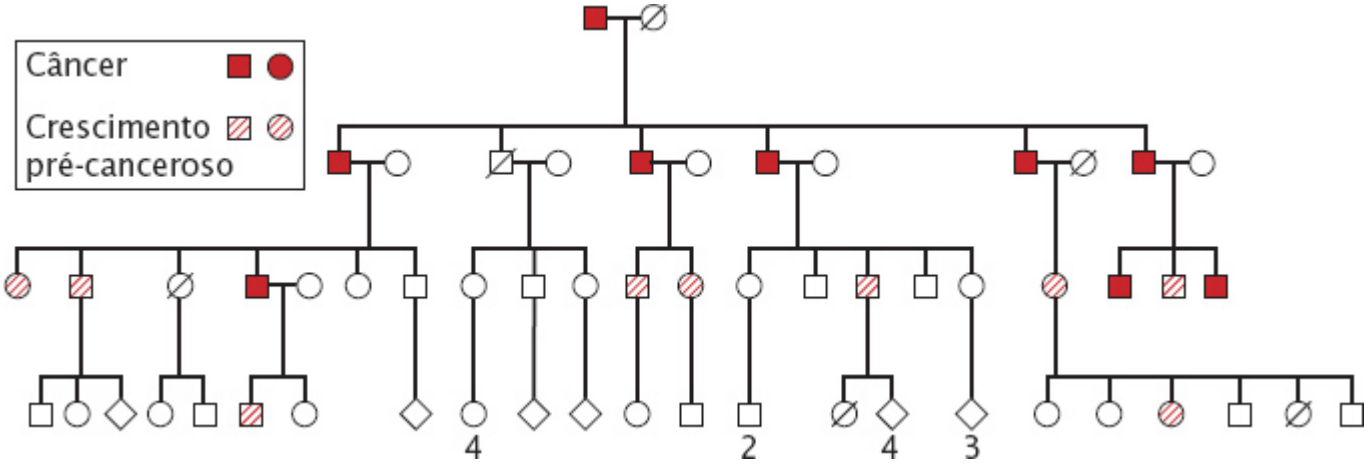


Figura 23.1 O câncer pancreático é herdado como um traço autossômico dominante em uma família que tem um gene *palladin* mutante. (De K. L. Pogue et al., Plos Medicine 3:2216-2228, 2006.)

A descoberta do elo do gene *palladin* com o câncer pancreático ilustra o poder da genética molecular moderna para solucionar a natureza biológica do câncer. Neste capítulo, examinamos a natureza genética do câncer, uma doença que é fundamentalmente genética, mas não é herdada. Começamos ao analisar a natureza do câncer e como as múltiplas alterações genéticas são necessárias para transformar uma célula normal em uma célula cancerosa. Então analisamos alguns dos tipos de genes que contribuem com o câncer, incluindo os oncogenes e genes supressores de tumor, genes que controlam o ciclo celular, genes que codificam proteínas de sistemas de reparo do DNA e telomerase, e genes que, como *palladin*, ajudam a disseminar o câncer. A seguir, vamos examinar as mudanças epigenéticas associadas com o câncer e como os genes específicos contribuem para o progresso do câncer do cólon. Finalmente, vamos discutir as mutações cromossômicas associadas com o câncer e o papel dos vírus em alguns cânceres.

23.1 O câncer é um grupo de doenças caracterizado pela proliferação celular

Uma em cada cinco pessoas nos EUA morrerá de câncer e os tratamentos para câncer custam bilhões de dólares por ano. O câncer não é uma doença simples; na verdade, é um grupo heterogêneo de distúrbios caracterizado por células que não respondem aos controles normais de divisão. As células do câncer se dividem rápida e continuamente, criando tumores que comprimem as células normais e acabam roubando os nutrientes de tecidos saudáveis (**Figura 23.2**). As células de um tumor avançado podem se separar do tumor e se deslocar para locais distantes no corpo, onde se implantam e desenvolvem novos tumores. Os cânceres mais comuns nos EUA são de próstata, mama, pulmão, cólon e reto e sangue (**Quadro 23.1**).

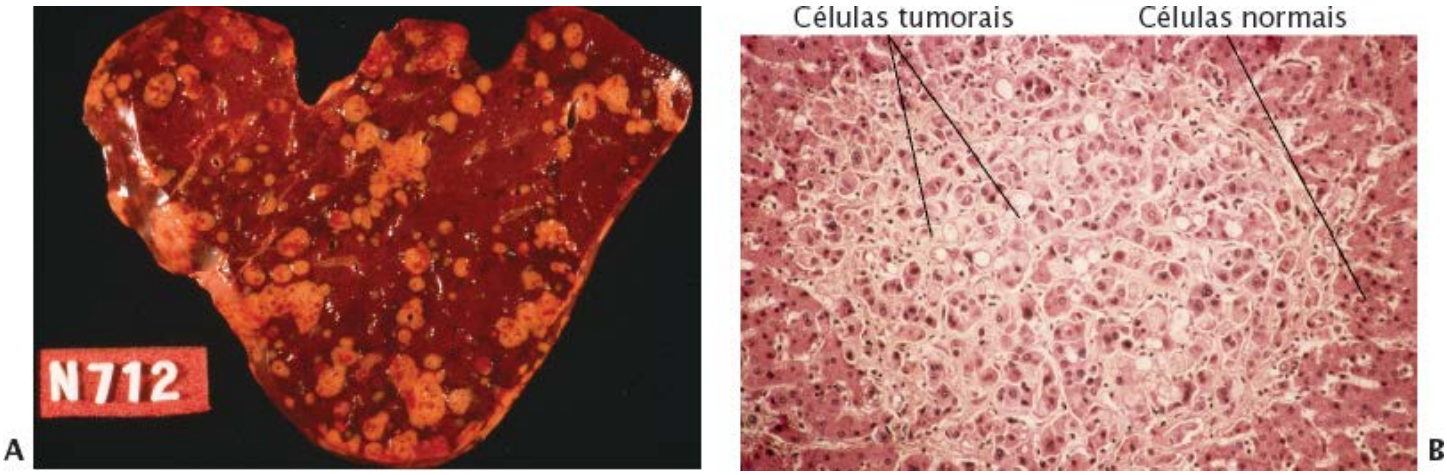


Figura 23.2 A proliferação anormal de células do câncer produz um tumor que comprime as células normais. A. As massas metastáticas de câncer de mama (*protrusões em branco*) crescendo em um fígado humano. **B.** Uma micrografia óptica de um corte de fígado com tumores. As células do câncer são claras, desbotadas; as células mais escuras são hepatócitos saudáveis. (CNRI/Science Source.)

Quadro 23.1 Incidências estimadas de vários cânceres e mortes causadas pelo câncer nos EUA em 2013.

Tipo de câncer	Novos casos por ano	Mortes por ano
Próstata	238.590	29.720
Mama	234.580	40.030
Pulmão e brônquio	228.190	159.480
Cólon e reto	142.820	50.830
Linfoma	79.030	20.200
Melanoma	76.690	9.480
Bexiga	72.570	15.210
Útero	49.560	8.190
Leucemias	48.610	23.720
Pâncreas	45.220	38.460
Cavidade oral e faringe	41.380	7.890
Fígado	30.640	21.670
Cérebro e sistema nervoso	23.130	14.080
Ovário	22.240	14.030
Estômago	21.600	10.990
Colo uterino	12.340	4.030
Cânceres de tecidos moles, incluindo o coração	11.410	4.390
Todos os cânceres	1.660.290	580.350

Fonte: American Cancer Society, Cancer Facts and Figures, 2013 (Atlanta: American Cancer Society, 2013), p. 4.

Formação do tumor

As células normais crescem, se dividem, amadurecem e morrem em resposta a um complexo conjunto de sinais internos e externos. Uma célula normal recebe sinais estimuladores e inibitórios e seu crescimento e sua divisão são regulados por um delicado equilíbrio entre essas forças opostas. Em uma célula do câncer, um ou mais sinais são comprometidos, o que faz com que a célula prolifere em uma taxa absurdamente alta. À medida que perdem sua resposta aos controles normais, essas células gradualmente perdem seu formato e limites regulares, consequentemente formando uma massa distinta de células anormais, um tumor. Se as células do tumor permanecem localizadas, o tumor é considerado benigno, se as células invadem outros tecidos, o tumor é considerado **maligno**. As células que se deslocam para outros locais no corpo, onde estabelecem tumores secundários, sofreram **metástase**.

Câncer como uma doença genética

O câncer resulta de defeitos fundamentais na regulação da divisão celular e seu estudo é importante não apenas para a saúde pública, mas também para nossa compreensão sobre a biologia celular. Ao longo dos anos, foram desenvolvidas muitas ideias para explicar o câncer, mas agora sabemos que a maioria dos, se não todos, cânceres surgem a partir de defeitos no DNA.

Evidências genéticas de câncer. As observações iniciais sugeriram que o câncer poderia resultar de um dano genético. Primeiro, muitos agentes, como a radiação ionizante e substâncias químicas, que provocam mutação, também causam câncer (são carcinógenos, ver *Capítulo 18*). Segundo, alguns cânceres estão associados com anomalias cromossômicas específicas. Cerca de 90% das pessoas com leucemia mieloide crônica, por exemplo, têm uma translocação recíproca entre o cromossomo 22 e o cromossomo 9. Terceiro, alguns tipos específicos de cânceres tendem a ocorrer em determinadas famílias. O

retinoblastoma, um raro câncer de infância da retina, aparece com alta frequência em algumas famílias e é herdado como um traço autossômico dominante, sugerindo que um único gene é responsável por esses casos da doença.

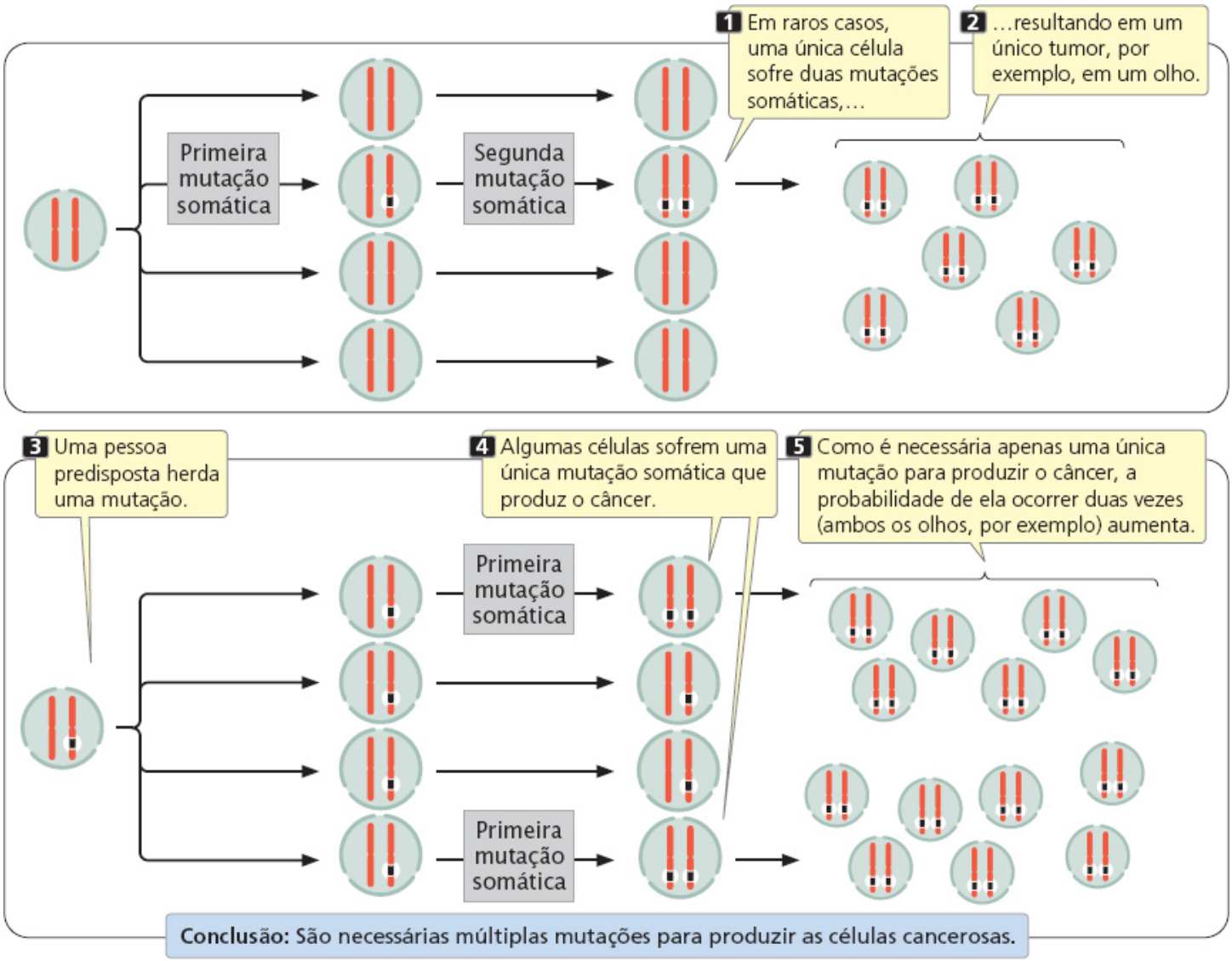


Figura 23.3 Alfred Knudson propôs que o retinoblastoma é o resultado de dois defeitos genéticos separados, ambos necessários para o câncer se desenvolver.

Embora essas observações tenham sugerido que os genes participem do câncer, a teoria do câncer como uma doença genética tem vários problemas relevantes. Se o câncer for hereditário, todas as células no corpo devem receber o gene causador do câncer e, portanto, todas as células devem se tornar cancerosas. Nos tipos de cânceres que ocorrem nas famílias, entretanto, os tumores aparecem apenas em alguns tecidos e apenas quando a pessoa alcança uma idade avançada. Finalmente, muitos cânceres não apresentam caráter familiar e mesmo com relação a esses cânceres que ocorrem, surgem casos isolados em famílias sem histórico da doença.

Modelo de multietapas de Knudson para o câncer. Em 1971, Alfred Knudson propôs um modelo para explicar a base genética do câncer. Ele estava estudando o retinoblastoma, que em geral é unilateral, mas, em alguns casos, surge em ambos os olhos. Knudson descobriu que, quando o retinoblastoma aparecia em ambos os olhos, isso ocorria em idade precoce e afetava crianças cujos parentes próximos também tinham a doença.

Ele propôs que o retinoblastoma resulta de dois defeitos genéticos separados, ambos necessários para o câncer se desenvolver (**Figura 23.3**). E sugeriu que, nos casos de acometimento unilateral, uma única célula em um olho sofre duas mutações recessivas. Como a chance de essas duas mutações ocorrerem em uma única célula é remota, o retinoblastoma é raro e tipicamente se desenvolve em apenas um olho. Para o acometimento bilateral, Knudson propôs que a criança herda uma das duas mutações necessárias para o câncer, e cada célula tem essa mutação inicial. Nesses casos, é necessário que uma célula do olho sofra a segunda mutação para que o câncer surja. Como cada olho tem milhões de células, a probabilidade de que a segunda mutação ocorrerá em pelo menos uma célula de cada olho é alta, produzindo tumores bilaterais precocemente.

A proposta de Knudson sugere que o câncer é o resultado de um processo de várias etapas que requer várias mutações. Se

uma ou mais das mutações necessárias for herdada, são necessárias menos mutações adicionais para produzir o câncer e esse tenderá a percorrer nas famílias. A ideia de Knudson foi chamada de “hipótese dos dois eventos” porque, para o retinoblastoma, são necessárias apenas duas mutações para causar o tumor. Entretanto, para a maioria dos cânceres estão envolvidas mais de duas mutações na transformação de células normais em células cancerosas. No caso do retinoblastoma, as duas mutações necessárias ocorrem no mesmo *locus* (ambos os alelos sofrem mutação), mas para a maioria dos cânceres são necessárias mutações em diferentes *loci* para o desenvolvimento da doença. A ideia de que o câncer é o resultado de múltiplas mutações tornou-se correta para a maioria dos cânceres.

A teoria genética de Knudson para o câncer foi confirmada pela identificação de genes, que quando sofrem mutação, causam o câncer. Atualmente, reconhecemos que o câncer é fundamentalmente uma doença genética, embora alguns cânceres sejam herdados. A maioria dos tumores surge a partir de mutações somáticas que se acumulam na vida de uma pessoa, por meio de mutação espontânea ou em resposta a mutágenos ambientais. **Resolva o Problema 24**

Evolução clonal dos tumores. O câncer começa quando uma única célula sofre uma mutação que faz com que ela se divida em uma taxa rápida e anormal. A célula prolifera, dando origem a um clone de células, cada qual carreia a mesma mutação. Como as células do clone se dividem mais rapidamente que o normal, elas se sobrepõem às outras células. Uma mutação adicional que surge em algumas das células do clone pode acentuar a capacidade dessas células em proliferar, e as células com ambas as mutações se tornam as células mais comuns no clone. Consequentemente, elas podem ser ultrapassadas por células com mais mutações que acentuam a proliferação. Nesse processo, chamado de **evolução clonal**, as células adquirem mais mutações que permitem que elas se tornem cada vez mais agressivas nas suas propriedades proliferativas (**Figura 23.4**).

A taxa da evolução clonal depende da frequência na qual surgem novas mutações. Qualquer defeito genético que permita que surjam mais mutações vai acelerar o progresso do câncer. Foi descoberto que os genes que regulam o reparo do DNA sofreram mutação nas células de cânceres avançados, e os distúrbios herdados do reparo do DNA são caracterizados por incidências crescentes de câncer. Como os mecanismos de reparo do DNA normalmente eliminam muitas das mutações que surgem, as células com sistemas de reparo do DNA com defeito estão mais propensas a reter as mutações do que as células normais, incluindo as mutações nos genes que regulam a divisão celular. O xeroderma pigmentoso, por exemplo, é um raro distúrbio causado por um defeito no reparo do DNA (ver *Capítulo 18*). As pessoas com essa doença apresentam taxas elevadas de câncer de pele quando expostas à luz do sol (que induz a mutação). Da mesma forma, o câncer de mama pode ser causado por mutações em *BRCA1* e *BRCA2*, dois genes que atuam no reparo do DNA.

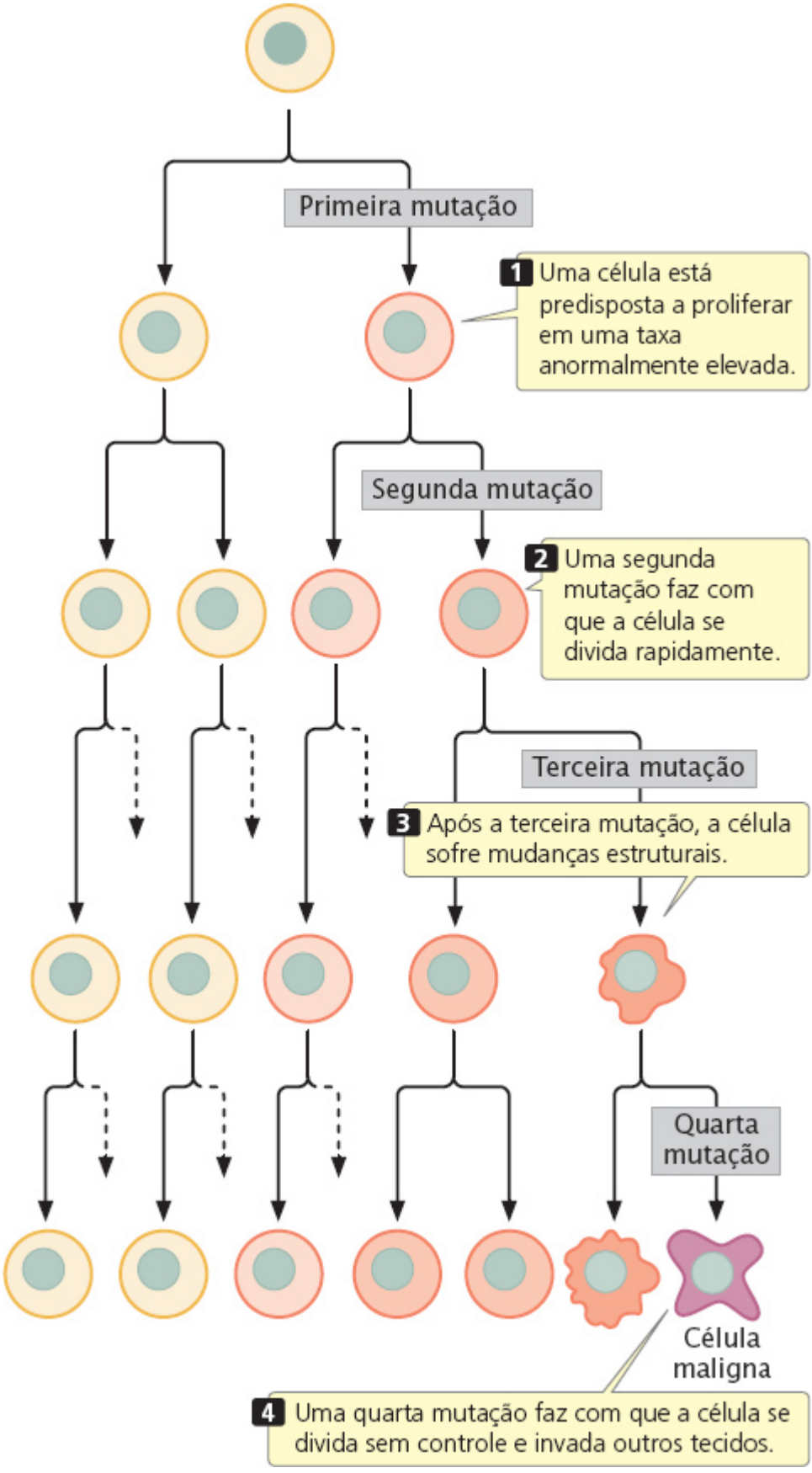


Figura 23.4 Pela evolução clonal, as células tumorais adquirem múltiplas mutações que permitem que se tornem cada vez mais agressivas e proliferativas. Para preservar o espaço, é usada uma seta pontilhada para representar uma segunda célula do mesmo tipo em cada caso.

As mutações em genes que afetam a segregação cromossômica também podem contribuir com a evolução clonal dos tumores. Muitas células de câncer são aneuploides (têm cópias extras ou não de cromossomos individuais, ver *Capítulo 8*) e, é claro, mutações cromossômicas contribuem para o progresso do câncer ao duplicar alguns genes (os genes nos cromossomos extras) e eliminar outros (nos cromossomos eliminados). Defeitos celulares que interferem na separação dos cromossomos aumentam a aneuploidia e podem acelerar o progresso do câncer.

Conceitos

O câncer é fundamentalmente uma doença genética. Em geral, são necessárias mutações em vários genes para produzir o câncer. Se uma dessas mutações for herdada, são necessárias menos mutações somáticas para o câncer surgir, e a pessoa pode ter predisposição para o câncer. A evolução clonal é o acúmulo de mutações em um clone de células.

Checagem dos conceitos 1

Como o modelo de várias etapas do câncer explica a observação de casos esporádicos de retinoblastoma serem unilaterais, enquanto as formas hereditárias do câncer são bilterais?

Papel dos fatores ambientais no câncer

Embora o câncer seja uma doença genética, a maioria dos cânceres não é herdada, e muitos são influenciados pelos fatores ambientais. O papel dos fatores ambientais no câncer é sugerido por diferenças na incidência de cânceres específicos pelo mundo (**Quadro 23.2**). Os resultados dos estudos mostram que populações migratórias incorporam a incidência de câncer de seu país de chegada. Por exemplo, as taxas de câncer em geral são consideravelmente mais baixas no Japão do que no Havaí. Entretanto, em uma única geração após a migração para o Havaí, os japoneses desenvolveram cânceres em taxas semelhantes às dos havaianos nativos. O câncer crescente entre os imigrantes se deve ao fato de que eles são expostos aos mesmos fatores ambientais que os nativos.

Vários fatores ambientais contribuem para o câncer, mas os que têm os maiores efeitos incluem fumo, dieta, obesidade, álcool e radiação UV (**Quadro 23.3**). Outros fatores ambientais que induzem o câncer são alguns tipos de substâncias químicas, como benzeno (usado como solvente industrial), benzo[a]pireno (encontrado no cigarro) e bifenis policlorados (PCBs, usados em transformadores industriais e capacitores). A maioria dos fatores ambientais associados com câncer causa mutações somáticas que estimulam a divisão celular ou afetam de outra forma o processo do progresso do câncer.

Quadro 23.2 Exemplos de variação geográfica na incidência de câncer.

Tipo de câncer	Localização	Taxa de incidência*
Lábio	Canadá (Newfoundland)	15,1
	Brasil (Fortaleza)	1,2
Nasofaringe	Hong Kong	30,0
	EUA (Utah)	0,5
Cólon	EUA (Iowa)	30,1
	Índia (Mumbai)	3,4
Pulmão	EUA (New Orleans, afro-americanos)	110,0
	Costa Rica	17,8
Próstata	EUA (Utah)	70,2
	China (Xangai)	1,8
Bexiga	EUA (Connecticut, brancos)	25,2
	Filipinas (Rizal)	2,8
Todos os cânceres	Suíça (Basileia)	383,3
	Kuwait	76,3

Fonte: C. Muir et al., Cancer Incidence in Five Continents, vol. 5 (Lyon: International Agency for Research on Cancer, 1987), Table 12-2.

*A taxa de incidência é a taxa padronizada por idade nos homens por 100.000 da população.

Quadro 23.3 Porcentagem de casos de câncer no Reino Unido causados por fatores ambientais.

Fator	Porcentagem de casos de câncer
Fumo	19,4
Dieta	9,2
Sobrepeso e obesidade	5,5
Álcool etílico	4,0
Profissão	3,7
Radiação (UV)	3,5
Infecções	3,1
Radiação (ionizante)	1,8
Todos os fatores ambientais	42,7

Fonte: Parkin, D.M., L. Boyd, and L.C. Walker. 2011. Fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. British Journal of Cancer 105:S77-S81.

Fatores ambientais podem interagir com predisposições genéticas ao câncer. Por exemplo, o câncer de pulmão está associado com tabagismo, um fator ambiental. Os estudos de associação genômica ampla (ver *Capítulo 20*) revelaram que a variação em vários genes predispõe algumas pessoas ao câncer de pulmão induzido pelo fumo. As variantes em alguns desses genes deixam as pessoas mais predispostas a se viciarem ao fumo. Outros genes que predispõem codificam receptores que se ligam a potenciais carcinógenos presentes no cigarro. **Resolva o Problema 23**

23.2 Mutações em vários tipos diferentes de genes contribuem para o câncer

Como já aprendemos, o câncer é uma doença causada por alterações no DNA. Entretanto, existem muitos tipos diferentes de alterações genéticas que podem contribuir para o câncer. Já foram identificados mais de 350 genes humanos diferentes que contribuem com o câncer; o número real deve ser muito maior. Pesquisa em camundongos sugere que mais de 2.000 genes podem, quando sofrem mutação, contribuir com o desenvolvimento do câncer. Nas seções a seguir, vamos analisar alguns dos tipos diferentes de genes que participam do câncer.

Oncogenes e genes supressores de tumor

Existem dois tipos básicos de sinais que regulam a divisão celular: as moléculas que estimulam a divisão celular e as que a inibem. Esses mecanismos de controle são semelhantes ao acelerador e ao freio de um carro. Nas células normais (mas, esperamos, não no seu carro), os aceleradores e os freios são acionados ao mesmo tempo, fazendo com que a divisão celular prossiga a uma velocidade adequada.

Como a divisão celular é afetada pelos aceleradores e freios, o câncer pode surgir a partir de mutações no tipo de sinal e existem várias rotas essencialmente diferentes para o câncer. Um gene estimulador pode se tornar hiperativo ou ativo em momentos inadequados, análogo a ter um acelerador de carro pisando fundo. As mutações nos genes estimuladores são dominantes porque até a quantidade reduzida do produto do gene fabricado por um único alelo é suficiente para promover o efeito estimulante. Os genes estimuladores mutantes de ação dominante responsáveis pelo câncer são chamados de **oncogenes** (**Figura 23.5 A**). A divisão celular também pode ser estimulada quando os genes inibitórios são *inativados*, análogo a ter um freio com defeito no carro. Os genes inibitórios mutantes em geral têm efeitos recessivos; assim, ambas as cópias devem ser mutantes para remover toda a inibição. Os genes inibitórios no câncer são chamados de **genes supressores de tumor** (**Figura 23.5 B**). Muitas células de câncer têm mutações nos oncogenes e nos genes supressores de tumor.

Embora os oncogenes ou os genes supressores de tumor mutantes ou ambos sejam necessários para produzir o câncer, as mutações nos genes de reparo do DNA podem aumentar a probabilidade de adquirir mutações nesses genes. Ter genes de reparo do DNA mutantes é como ter um mecânico de carro incompetente que não faz os reparos necessários em um acelerador ou freio quebrado.

Oncogenes. Os oncogenes foram os primeiros genes causadores de câncer a serem identificados. Em 1909, um fazendeiro levou ao médico Peytin Rous uma galinha com um grande tumor de tecido conjuntivo (sarcoma) crescendo em seu peito.

Quando Rous injetou pedaços desse tumor em outras galinhas, elas também desenvolveram sarcomas. Ele fez experimentos que demonstraram que os tumores eram transmitidos por um vírus, que ficou conhecido como o vírus do sarcoma de Rous. Dessa forma, vários outros vírus causadores de câncer foram isolados de vários tecidos animais. Acreditava-se que esses vírus carreassem um gene causador de câncer que era transferido para a célula hospedeira. O primeiro oncogene, chamado *src*, foi isolado do vírus do sarcoma de Rous em 1970.

Em 1975, Michael Bishop, Harold Varmus e colaboradores começaram a usar sondas para oncogenes virais para pesquisar sequências relacionadas nas células normais. Eles descobriram que os genomas de todas as células normais carreavam sequências de DNA intimamente relacionadas com os oncogenes. Esses genes das células normais eram chamados de **proto-oncogenes**. Eles são responsáveis pelas funções celulares básicas nas células normais, mas quando são mutantes, se tornam oncogenes que contribuem para o desenvolvimento do câncer. Quando um vírus infecta uma célula, um proto-oncogene pode ser incorporado ao genoma viral pela recombinação. Dentro do genoma viral, o proto-oncogene pode sofrer mutação para um oncogene, que, quando inserido de volta à célula, leva a rápida divisão celular e ao câncer. Como é mais provável que os proto-oncogenes sofram mutação ou recombinação dentro de um vírus, a infecção viral está associada com o câncer.

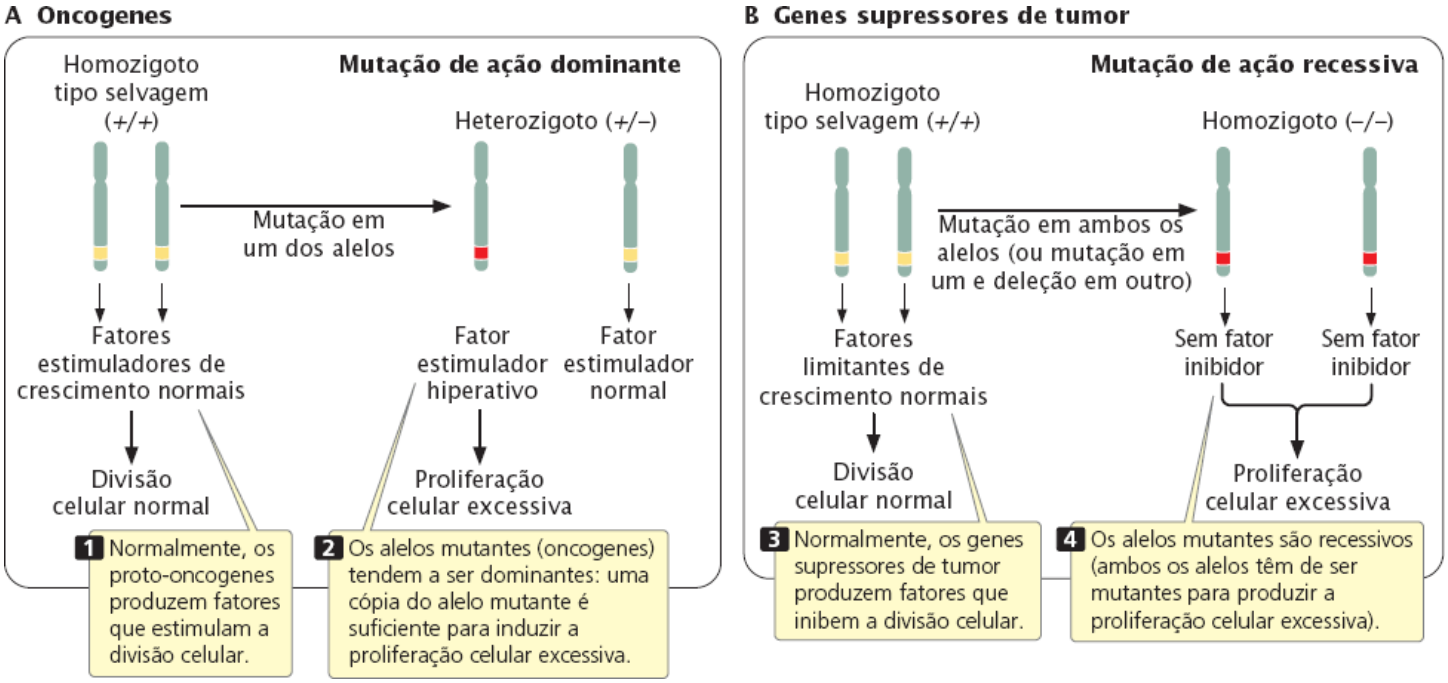


Figura 23.5 Oncogenes e genes supressores de tumor contribuem para o câncer, mas seus modos de ação e dominância são diferentes.

Os proto-oncogenes podem ser convertidos em oncogenes nos vírus de várias formas diferentes. A sequência de proto-oncogene pode ser alterada ou truncada à medida que é incorporada ao genoma viral. Essa cópia mutante do gene pode então produzir uma proteína alterada que provoca a proliferação celular descontrolada. Por outro lado, pela recombinação, um proto-oncogene pode ficar próximo de um promotor ou acentuador viral, fazendo com que o gene seja superexpresso. Finalmente, em alguns casos, a função de um proto-oncogene na célula hospedeira pode ser alterada quando um vírus insere seu próprio DNA no gene, interrompendo sua função normal. Enquanto os vírus são capazes de converter proto-oncogenes em oncogenes, a maioria dos proto-oncogenes sofre mutação para formar oncogenes sem o envolvimento de um vírus.

Foram identificados muitos oncogenes em experimentos nos quais fragmentos selecionados de DNA são adicionados a células em cultura. Algumas células captam o DNA, e, se essas células se tornarem cancerosas, então o fragmento de DNA adicionado à cultura deve ter um oncogene. Os fragmentos podem então ser sequenciados e o oncogene, identificado. Já foram descobertos vários oncogenes (**Quadro 23.4**). Acredita-se que cerca de 90% de todos os genes de câncer sejam oncogenes dominantes.

Quadro 23.4 Alguns oncogenes e funções de seus proto-oncogenes correspondentes.

Gene	Função normal	Câncer no qual o gene é mutante
erbB	Parte do receptor do fator de crescimento	Muitos tipos de câncer

fos	Fator de transcrição	Osteossarcoma e carcinoma endometrial
jun	Fator de transcrição, controle do ciclo celular	Câncer de pulmão, câncer de mama
myc	Fator de transcrição	Linfomas, leucemias, neuroblastoma
ras	Ligação a GTP e GTPase	Muitos tipos de câncer
sis	Fator de crescimento	Glioblastomas e outros cânceres
src	Proteína tirosinoquinase	Muitos tipos de câncer

Genes supressores de tumor. É mais difícil identificar os genes supressores de tumor do que os oncogenes porque eles *inibem* o câncer e são recessivos, ambos alelos precisam sofrer mutação antes que a inibição da divisão celular seja removida. Como a *falha* da sua função promove a proliferação da célula, os genes supressores de tumor não podem ser identificados ao adicioná-los às células e procurar por câncer. Acredita-se que cerca de 10% de todos os genes responsáveis pelo câncer sejam genes supressores de tumor.

Em geral, são necessários defeitos em ambas as cópias de um gene supressor de tumor para gerar o câncer; um organismo pode herdar uma cópia defeituosa desse gene (é heterozigoto para a mutação que causa câncer) e não ter câncer porque o alelo normal remanescente fabrica o produto supressor do tumor. Entretanto, esses heterozigotos estão predispostos ao câncer porque apenas a inativação ou perda de um alelo remanescente é necessária para eliminar completamente o produto supressor de tumor. A inativação do alelo do tipo selvagem remanescente nos heterozigotos é chamada de **perda da heterozigosidase**. Um mecanismo comum para a perda da heterozigosidase é a deleção no cromossomo que carreava a cópia normal do gene supressor do tumor (**Figura 23.6**).

Entre os primeiros genes supressores de tumor a serem identificados estava o gene responsável pelo retinoblastoma. Em 1985, Raymond White e Webster Cavenne mostraram que grandes segmentos do cromossomo 13 estavam ausentes nas células dos tumores do retinoblastoma, e posteriormente, foi isolado o gene supressor de tumor a partir desses segmentos. Outro exemplo de um gene supressor de tumor é o *BRCA1*; as suas mutações estão associadas com maior risco de cânceres de mama e ovário. O *BRCA1* produz uma proteína que normalmente ajuda no reparo de quebras de fita dupla no DNA por recombinação homóloga (ver *Capítulo 18*). Ele também atua como um fator de transcrição e interage com as enzimas histona desacetilases, que afetam a transcrição. Já foram descobertos vários genes supressores de tumor (**Quadro 23.5**).

Às vezes, a mutação ou a perda de um único alelo de um gene recessivo supressor de tumor é suficiente para desenvolver o câncer. Esse efeito, o surgimento de um traço em uma célula ou organismo que é heterozigoto para um traço normalmente recessivo, é chamado de **haploinsuficiência**. Acredita-se que esse fenômeno ocorra por causa dos efeitos de dosagem: o heterozigoto produz apenas metade do produto codificado pelo gene supressor do tumor. Normalmente, isso é suficiente para os processos celulares evitarem a formação de tumor, mas é menos que a quantidade ideal, e outros fatores podem combinar-se com esse produto supressor de tumor reduzido para desenvolver o câncer. **Resolva o Problema 25**

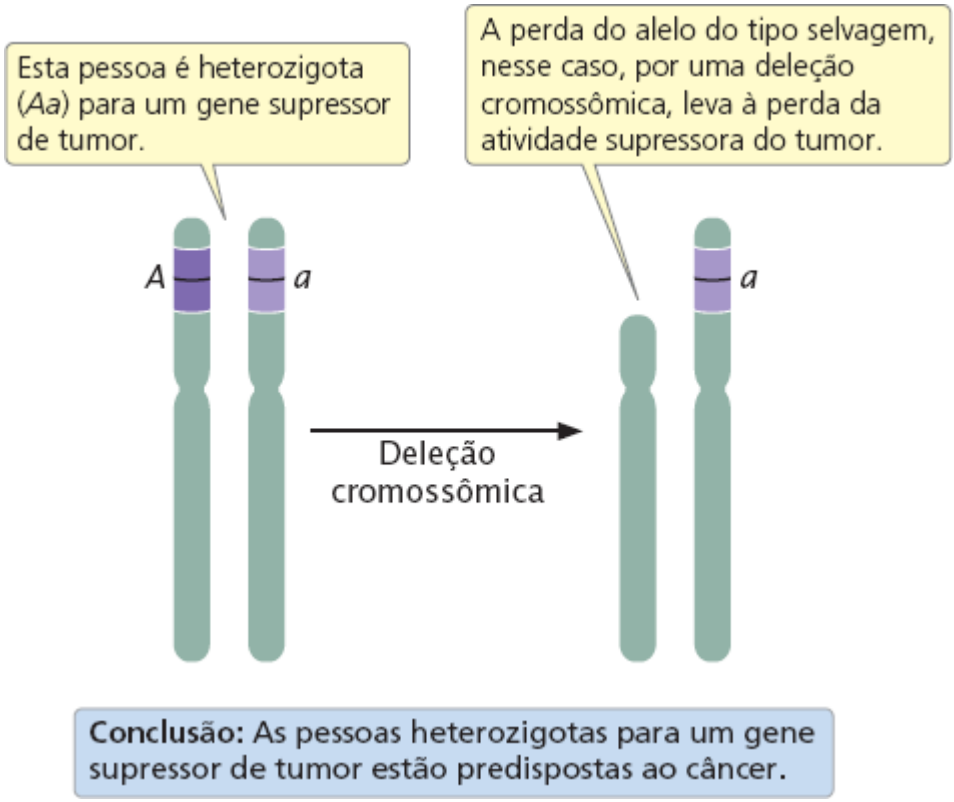


Figura 23.6 **A perda da heterozigosidase produz câncer em uma pessoa heterozigota para um gene supressor de câncer.**

Conceitos

Os proto-oncogenes são genes que controlam as funções celulares normais; quando sofrem mutação, se tornam oncogenes que estimulam a proliferação celular. Eles tendem a ter ação dominante. Normalmente, os genes supressores de tumor inibem a proliferação celular; quando sofrem mutação, permitem a proliferação celular. Os genes supressores de tumor podem ter ação recessiva. Organismos heterozigotos para os genes supressores de tumor são, com frequência, predispostos para o câncer.

Checagem dos conceitos 2

Por que os oncogenes têm ação dominante, enquanto os genes supressores de tumor têm ação recessiva?

Quadro 23.5 Alguns genes supressores de tumor e suas funções normais.

Gene	Função normal	Câncer no qual o gene é mutante
APC	Proteína scaffold, interage com microtúbulos	Colorretal
BRCA1	Reparo do DNA, fator de transcrição	Mama e ovário
CDKN2A	Regula a divisão celular	Melanoma
NF1	Ativador da GTPase	Neurofibromatose
p53	Regula a divisão celular	Muitos tipos de câncer
RB	Regula a divisão celular	Retinoblastoma

Mutações nos genes que controlam o ciclo de divisão celular

O ciclo celular é o processo normal pelo qual as células sofrem crescimento e divisão. Normalmente, o progresso pelo ciclo celular é rigorosamente regulado de modo que as células se dividem apenas quando são necessárias mais células, quando todos

os componentes necessários para divisão estão presentes e quando o DNA foi replicado sem danos. Às vezes, entretanto, surgem erros em um ou mais dos componentes que regulam o ciclo celular. Esses erros fazem as células se dividirem em momentos ou taxas inadequados, levando ao câncer. Na verdade, muitos proto-oncogenes e genes supressores de tumor atuam normalmente ao ajudar a controlar o ciclo celular. Antes de analisar como os erros nesse sistema contribuem para o câncer, primeiro é preciso compreender como o ciclo celular é regulado.

Controle do ciclo celular. Como foi discutido no Capítulo 2, o ciclo celular é o período entre uma divisão celular e a próxima. As células que se dividem ativamente passam pelas fases G_1, S e G_2 da interfase e se movem diretamente para a fase M, quando ocorre a divisão celular. As células que não se dividem saem da fase G_1 para a fase G_0, na qual são funcionais, mas não crescem nem se dividem. O progresso de um estágio do ciclo celular para outro é influenciado por vários sinais internos e externos e é regulado por pontos-chave no ciclo chamados de pontos de verificação.

Por muitos anos, os eventos bioquímicos que controlam o progresso das células pelo ciclo celular eram totalmente desconhecidos, mas atualmente a pesquisa revelou muitos dos detalhes desse processo. Os eventos principais do ciclo celular são controlados por **quinases dependentes de ciclina** (CDKs) que são enzimas que adicionam grupos fosfato a outras proteínas. Às vezes a fosforilação ativa a outra proteína e às vezes a inativa. Como seu nome sugere, as CDKs são funcionais apenas quando se associam com outra proteína chamada **ciclina**. O nível de ciclina oscila no curso do ciclo celular; quando ligada a uma CDK, a ciclina especifica quais proteínas a CDK vai fosforilar. Cada ciclina aparece em um ponto específico no ciclo celular, em geral porque sua síntese e destruição são reguladas por outra ciclina. As ciclinas e CDKs são chamadas por diferentes nomes em diferentes organismos; neste livro usaremos os termos aplicados a essas moléculas nos mamíferos.

Transição G_1 para S. Vamos começar observando a transição G_1 para S. Como já foi mencionado, os pontos de verificação garantem que todos os componentes celulares estejam presentes e em bom estado de funcionamento antes que a célula prossiga para o próximo estágio do ciclo. O ponto de verificação G_1/S está em G_1, pouco antes de a célula entrar na fase S e replicar seu DNA. A célula é impedida de passar pelo ponto de verificação G_1/S pela proteína do retinoblastoma (RB) (**Figura 23.7**), que se liga a outra molécula chamada E2F e a mantém inativa. Em G_1, a ciclina D e a ciclina E aumentam sua concentração de forma contínua e se combinam com suas CDKs associadas. Ciclina-D-CDK e ciclina-E-CDK fosforilam as moléculas do RB. Posteriormente em G_1, a fosforilação de RB é completada, inativando-a. Sem os efeitos inibitórios de RB, a proteína E2F é liberada. A E2F é um fator de transcrição que estimula a transcrição dos genes que produzem as enzimas necessárias para a replicação do DNA e a célula segue para o estágio S do ciclo.

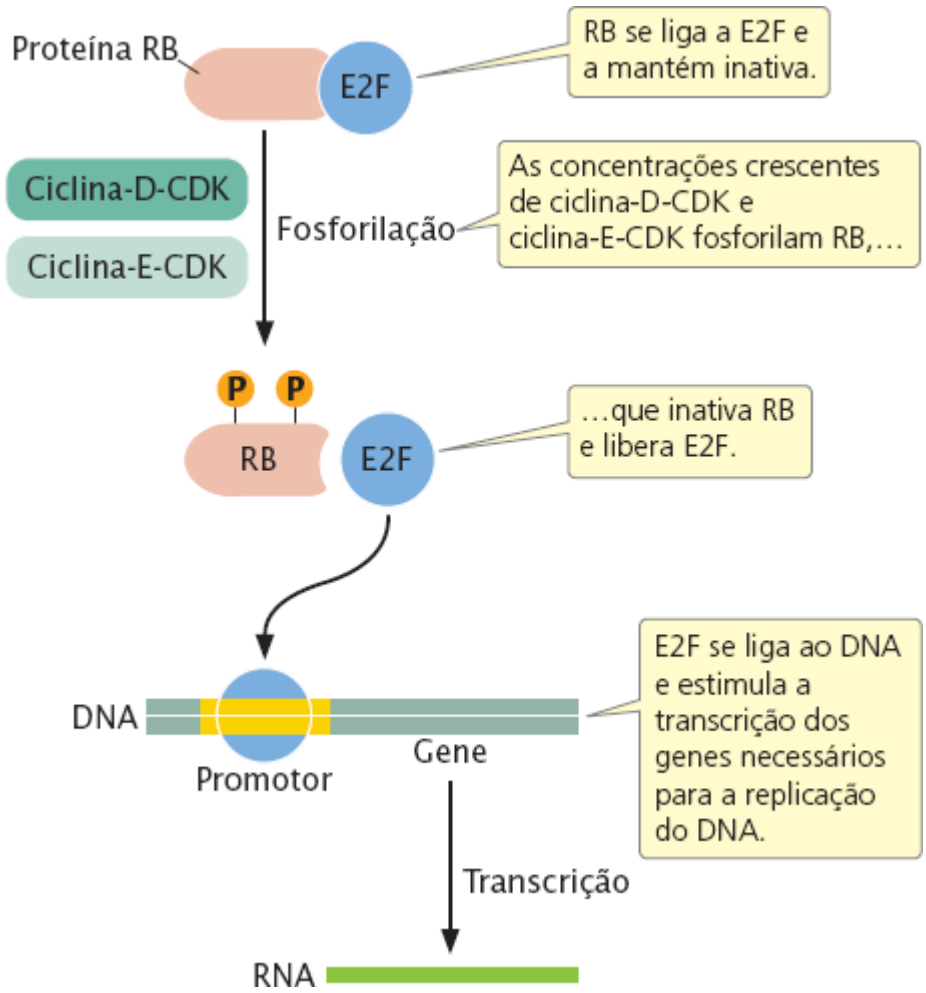


Figura 23.7 A proteína RB ajuda a controlar o progresso pelo ponto de verificação G_1/S ao se ligar ao fator de transcrição E2F.

Transição G_2 para M. A regulação da transição G_2 para M é semelhante à da transcrição G_1 para S. Na transição G_2 para M, a ciclina B se combina com CDK para formar um complexo inativo chamado *fator promotor da mitose* (MPF). Após o MPF ser formado, ele deve ser ativado com a remoção de um grupo fosfato (**Figura 23.8 A**). Durante G_1, os níveis de ciclina B estão reduzidos, então a quantidade de MPF também está reduzida. À medida que mais ciclina B é produzida, ela se combina com CDK para formar quantidades crescentes de MPF. Próximo do final de G_2, a quantidade de MPF ativo alcança um nível crítico, o que faz com que a célula se divida. A concentração de MPF continua a aumentar, alcançando um pico na mitose.

A forma ativa de MPF fosforila outras proteínas, que então desencadeiam muitos dos eventos associados com a mitose, como ruptura de membrana nuclear, formação do fuso e condensação do cromossomo. No final da metáfase, a ciclina B é abruptamente degradada, o que reduz a concentração de MPF e inicia a anáfase, ativa uma cadeia de eventos que por fim encerra a mitose (**Figura 23.8 B**). Em resumo, os níveis elevados de MPF ativo estimulam a mitose e os níveis reduzidos de MPF retornam para as condições de interfase.

O ponto de verificação G_2/M é no final de G_2, antes de a célula entrar na mitose. Vários fatores estimulam a síntese da ciclina B e a ativação do MPF, enquanto outros fatores inibem MPF. Juntos, esses fatores garantem que a mitose não seja iniciada até conseguir a condições adequadas para a divisão celular. Por exemplo, o dano ao DNA inibe a ativação do MPF; consequentemente, a célula para em G_2 e não sofre divisão.

Ponto de verificação da montagem do fuso. Outro ponto de verificação, chamado de ponto de verificação da montagem do fuso, está na metáfase. Esse ponto de verificação atrasa o início da anáfase até todos os cromossomos estarem alinhados na placa da metáfase e os cinetócoros-irmãos estarem presos às fibras do fuso a partir dos polos opostos. Se todos os cromossomos não estiverem adequadamente alinhados, o ponto de verificação bloqueia a destruição da ciclina B. A persistência da ciclina B mantém o MPF ativo e a célula em um estado mitótico. Um ponto de verificação adicional controla a saída da célula da mitose.

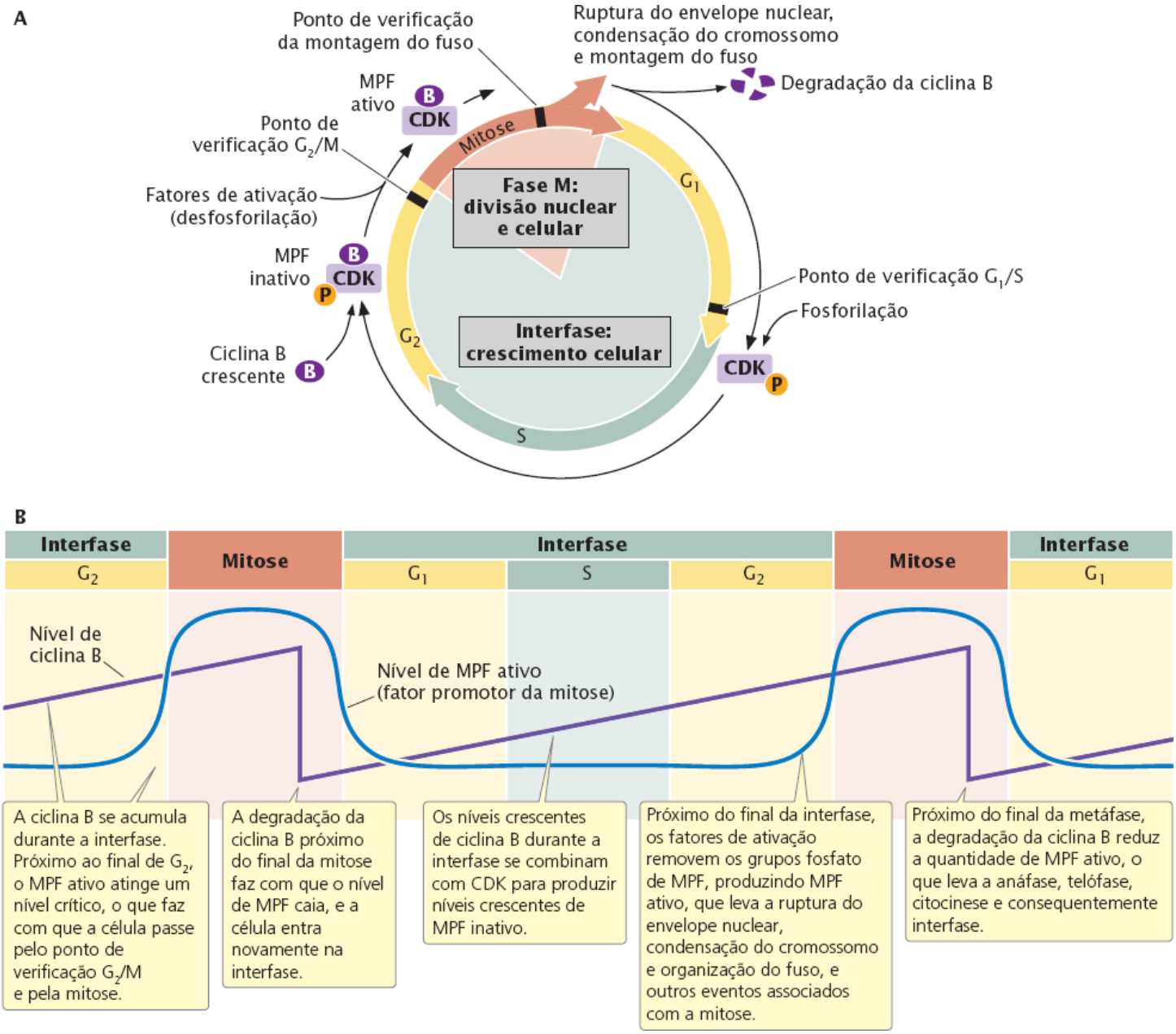


Figura 23.8 O progresso pelo ponto de verificação G_2/M é regulado pela ciclina B.

Mutações no controle do ciclo celular e câncer. Muitos cânceres são causados por defeitos no maquinário de regulação do ciclo celular. Por exemplo, as mutações no gene que codifica a proteína RB, que normalmente mantém a célula em G_1 até o DNA estar pronto para ser replicado, estão associadas com muitos cânceres, incluindo o retinoblastoma. Quando o gene *RB* sofre mutação, as células passam pelo ponto de verificação G_1/S sem os controles normais que evitam a proliferação celular. O gene que codifica a ciclina D (e estimula a passagem das células pelo ponto de verificação G_1/S) é expresso excessivamente em cerca de 50% dos cânceres de mama, assim como em alguns casos de câncer esofágico e de pele. Da mesma forma, o gene supressor de tumor *p53*, que sofre mutação em cerca de 75% de todos os cânceres de cólon, regula um potente inibidor da atividade da CDK.

Alguns proto-oncogenes e genes supressores de tumor participam da **apoptose**, um processo de morte celular programada no qual o DNA da célula é degradado, seu núcleo e citoplasma encolhem e a célula é fagocitada por outras células sem o vazamento de seu conteúdo. As células têm a capacidade de se autoavaliarem e, quando estão anormais ou danificadas, normalmente sofrem apoptose (ver *Capítulo 22*). As células do câncer apresentam mutações cromossômicas, dano ao DNA e outras anomalias celulares que normalmente estimulariam a apoptose e evitariam sua proliferação. Essas células têm mutações nos genes que regulam a apoptose e, portanto, não sofrem a morte celular programada. A capacidade de uma célula em iniciar a apoptose em resposta ao dano ao DNA, por exemplo, depende do *p53*, que está inativo em muitos cânceres humanos.

Conceitos

O progresso pelo ciclo celular é controlado em pontos de verificação, que são regulados por interações de

ciclinas e quinases dependentes de ciclina. Os genes que controlam o ciclo celular sofrem mutação nas células do câncer.

Checagem dos conceitos 3

Qual seria o efeito mais provável de uma mutação que faz com que a ciclina B seja incapaz de se ligar a CDK?

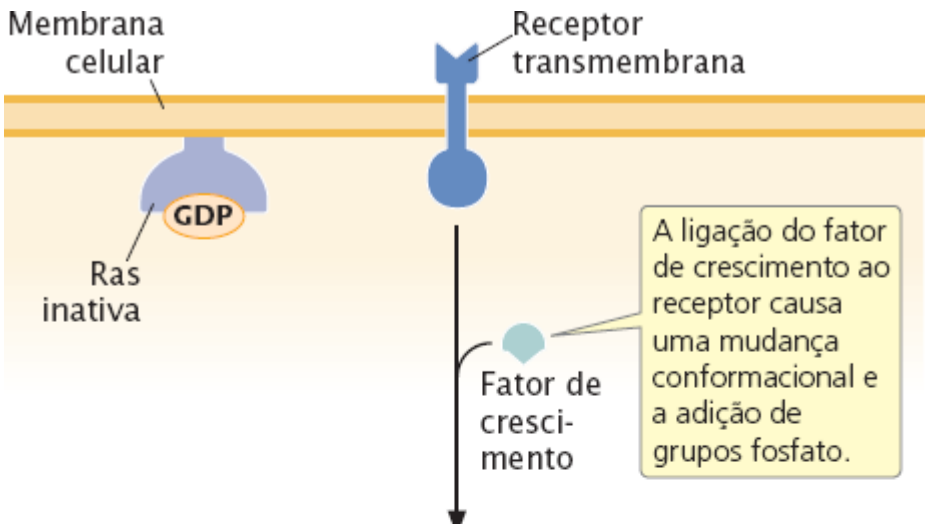
a. A célula passa pelo ponto de verificação G_2/M e entra na mitose mesmo quando o DNA não estiver replicado.
b. As células nunca passam pelo ponto de verificação G_1/S.
c. As células passam pela mitose mais rapidamente do que as células não mutantes.
d. As células não passam pelo ponto de verificação G_1/M e não entram em mitose.

Vias de transdução de sinal. Vários sinais internos e externos influenciam se a célula passa pelo ciclo celular e continua a se dividir. Os sinais externos são iniciados por hormônios e fatores de crescimento. Essas moléculas são incapazes de passar pela membrana celular por causa do seu tamanho e carga elétrica; elas exercem seus efeitos ao se ligarem a receptores na superfície da célula, que disparam uma série de reações intracelulares que carreiam a mensagem para o núcleo ou outro local dentro da célula. Esse tipo de sistema, no qual um sinal externo dispara uma cascata de reações intracelular que por fim produzem uma resposta específica, é chamado de **via de transdução de sinal**. Os defeitos nessas vias estão frequentemente associados com câncer.

Uma via de transdução de sinal começa com a ligação de uma molécula sinalizadora externa a um receptor específico que está incrustado na membrana celular. Os receptores nessas vias são compostos por três partes: (1) domínio extracelular que sai da célula e se liga à molécula sinalizadora; (2) um domínio transmembrana que atravessa a membrana e conduz o sinal para o interior da célula e (3) um domínio intracelular que se estende no citoplasma e, com a ligação da molécula sinalizadora, sofre uma mudança química ou conformacional que é transmitida para as moléculas da via de transdução de sinal no citoplasma. A ligação de uma molécula sinalizadora ao receptor ligado à membrana ativa uma proteína na via. Ao ser ativada, essa proteína ativa a molécula seguinte na via, ao adicionar ou remover grupos fosfatos ou provocando mudanças na conformação da proteína. A proteína recém-ativada ativa a molécula seguinte na via e, dessa forma, o sinal é transmitido por uma cascata de reações e, por fim, produz a resposta, estimulando ou inibindo o ciclo celular.

Na década passada, foram feitas muitas pesquisas para determinar as vias pelas quais os vários sinais influenciam o ciclo celular. Para ilustrar a transdução de sinal, vamos considerar a via de transdução de sinal Ras, que é importante no controle do ciclo celular. Cada proteína Ras alterna entre uma forma ativa e uma forma inativa. Na forma inativa, a proteína Ras se liga ao difosfato de guanosina (GDP); na forma ativa, se liga ao trifosfato de guanosina (GTP).

A via de transdução do sinal Ras é ativada quando um fator de crescimento como o fator de crescimento epidérmico (EGF) se liga a um receptor presente na membrana celular (**Figura 23.9**). A ligação do EGF provoca uma mudança conformacional no receptor e a adição de grupos fosfato a esse receptor. A adição de grupos fosfato permite que moléculas adaptadoras se liguem ao receptor. Essas moléculas adaptadoras se ligam ao receptor com uma molécula inativa da proteína Ras. As moléculas adaptadoras estimulam Ras a liberar GDP e se ligam ao GTP, ativando a Ras. A proteína Ras recém-ativada então se liga a uma forma inativa de outra proteína chamada Raf e a ativa. Após ativar Raf, Ras hidrolisa GTP a GDP, que converte Ras de volta para a forma inativa.



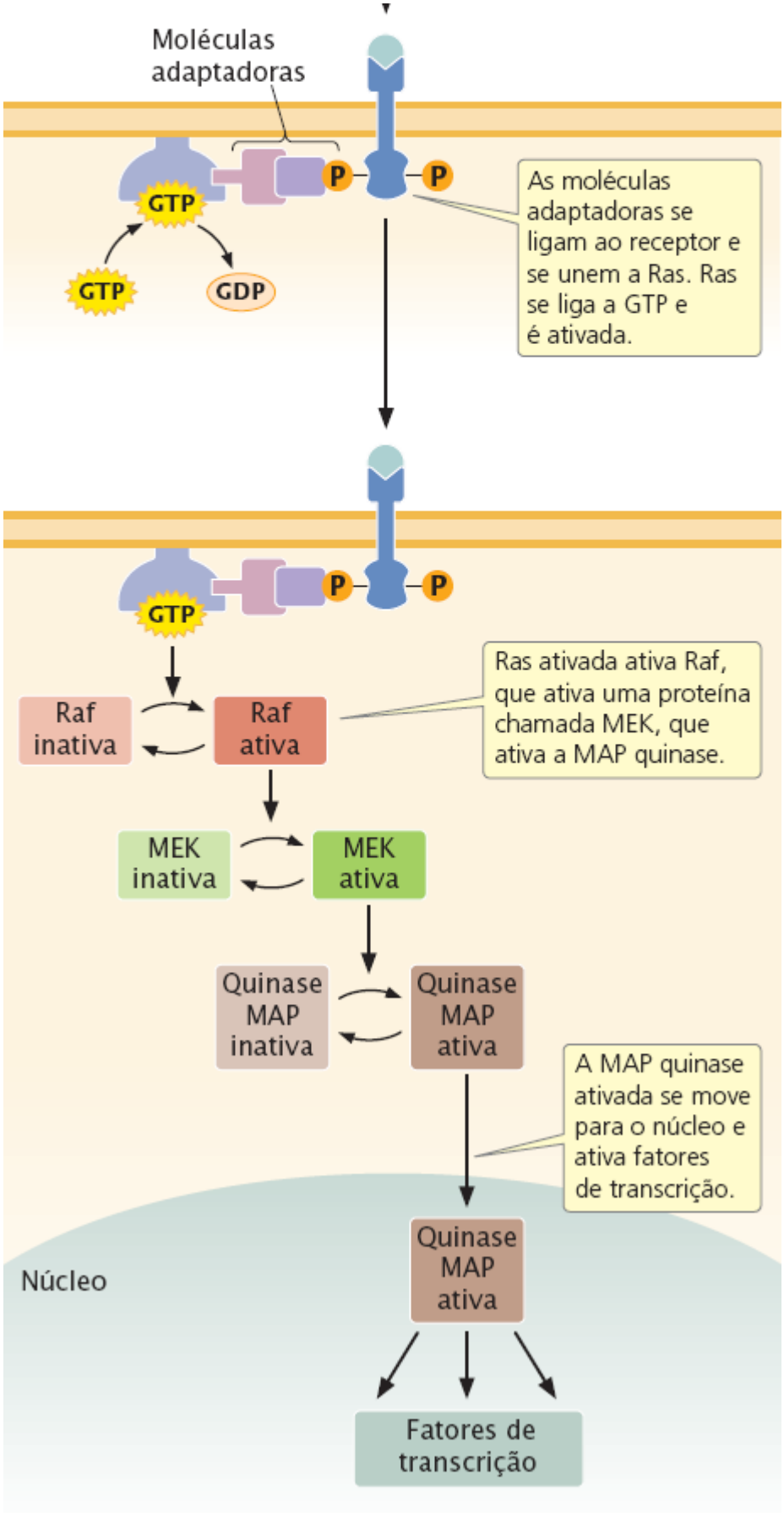


Figura 23.9 A via de transdução do sinal de Ras conduz sinais a partir de fatores de crescimento e hormônios para o núcleo e estimula o ciclo celular. As mutações nessa via contribuem para o câncer.

A Raf ativada inicia uma cascata de reações, terminando com a ativação de uma proteína chamada MAP quinase. A MAP quinase ativada se move para o núcleo e ativa vários fatores de transcrição que estimulam a transcrição de genes que participam do ciclo celular. Dessa forma, o sinal externo original promove a divisão celular. Foram identificadas várias outras vias de transdução que afetam o ciclo celular e a proliferação celular.

Como as vias de transdução de sinal ajudam a controlar o ciclo celular, defeitos nos seus componentes contribuem com o câncer. Por exemplo, os genes que codificam as proteínas Ras são frequentemente oncogenes, e as mutações nesses genes são encontradas em células de câncer: 75% dos tumores do pâncreas e 50% dos tumores na tireoide e cólon têm mutações nos genes *ras*. As mutações nesses genes produzem proteínas Ras mutantes que estão permanentemente ativadas e estimulam a divisão celular contínua.

Conceitos

Moléculas fora da célula produzem respostas intracelulares ao se ligar a um receptor de membrana e estimular uma cascata de reações intracelulares, conhecida como via de transdução de sinal. Muitas moléculas na via são proteínas que alternam entre as formas ativas e inativas. Os defeitos nessas vias estão frequentemente associados com câncer.

Checagem dos conceitos 4

As proteínas Ras são ativadas quando elas:

a. Ligam-se a GTP.
b. Liberam GTP.
c. Ligam-se a GDP.
d. Sofrem acetilação.

Genes de reparo do DNA

O câncer surge a partir do acúmulo de múltiplas mutações em uma única célula. Algumas células de câncer têm taxas normais de mutação e múltiplas mutações se acumulam porque cada mutação dá a célula uma vantagem replicativa sobre as células sem as mutações. Outras células do câncer podem ter taxas de mutação acima do normal em todos os seus genes, o que leva a uma mutação mais frequente de oncogenes e genes supressores de tumor. O que pode ser fonte dessas altas taxas de mutação em algumas células de câncer?

Dois processos controlam a taxa na qual as mutações surgem dentro de uma célula: (1) a taxa na qual os erros surgem durante e após o curso da replicação e (2) a eficiência com a qual esses erros são corrigidos. A taxa de erros na replicação é controlada pela fidelidade das DNA polimerases e outras proteínas no processo de replicação (ver *Capítulo 12*). Entretanto, defeitos nos genes que codificam as proteínas de replicação não estão fortemente ligados ao câncer.

A taxa de mutação é muito afetada por erros que são corrigidos pelos sistemas de reparo de DNA (ver *Capítulo 18*). Defeitos nos genes que codificam os componentes desses sistemas de reparo estão associados com vários cânceres. Pessoas com xeroderma pigmentoso, por exemplo, têm defeito no reparo de excisão de nucleotídios, um sistema de reparo celular importante que normalmente corrige o dano ao DNA causado por vários mutágenos, incluindo a luz ultravioleta. Da mesma forma, cerca de 13% dos cânceres colorretal, endometrial e gástrico têm células com defeito no sistema de reparo do pareamento errado, outro grande sistema de reparo na célula.

Um tipo específico de câncer de cólon chamado câncer colorretal não poliposo é herdado como um traço autossômico dominante. Nas famílias com essa doença, uma pessoa pode herdar um alelo mutante e um alelo normal de um gene que controla o reparo do pareamento errado. O alelo normal fornece níveis suficientes de proteína para que o reparo funcione, mas é muito provável que esse alelo se torne mutante ou seja perdido em pelo menos algumas células. Se isso acontecer, não há mais reparo de pareamento errado e essas células sofrem taxas de mutação acima do normal, levando a defeitos nos oncogenes e genes supressores de tumor que fazem a célula se proliferar.

Defeitos nos sistemas de reparo do DNA também podem contribuir para a geração de rearranjos cromossômicos e instabilidade genômica. Muitos sistemas de reparo de DNA fazem quebras de fita única e dupla no DNA durante o processo de reparo. Se essas quebras não forem feitas corretamente, surgem rearranjos cromossômicos.

Genes que regulam a telomerase

Outro fator que pode contribuir para o progresso do câncer é a ativação inadequada da enzima telomerase. Os telômeros são sequências especiais nas extremidades dos cromossomos eucarióticos. Lembre-se de que as extremidades dos cromossomos não podem ser replicadas e os telômeros ficam menores a cada divisão celular. Esse encurtamento leva a destruição do cromossomo e morte celular, então as células somáticas são capazes apenas de fazer um número limitado de divisões celulares.

Nas células germinativas e nas células-tronco, a telomerase replica as extremidades dos cromossomos (ver *Capítulo 12*), mantendo os telômeros, mas normalmente essa enzima não é expressa nas células somáticas. Em muitas células tumorais,

entretanto, as sequências que regulam a expressão do gene da telomerase sofrem mutação, permitindo que a enzima seja expressa e a célula é capaz de se dividir de forma ilimitada. Essa mutação permite que as células do câncer se dividam indefinidamente. Embora a expressão da telomerase pareça contribuir para o desenvolvimento de muitos cânceres, seu papel preciso no progresso do tumor não é conhecido e está sob investigação.

Genes que promovem a vascularização e a disseminação dos tumores

Um conjunto final de fatores que contribui para o progresso do câncer inclui genes que afetam o crescimento e a disseminação dos tumores. O oxigênio e os nutrientes, que são essenciais para a sobrevivência e crescimento dos tumores, são fornecidos pelos vasos sanguíneos e o crescimento de novos vasos (angiogênese) é importante para o progresso do tumor. A angiogênese é estimulada por fatores de crescimento e outras proteínas codificadas por genes cuja expressão é cuidadosamente regulada nas células normais. Nas células tumorais, os genes que codificam essas proteínas são superexpressos em comparação com as células normais e os inibidores dos fatores que promovem a angiogênese podem essar inativados ou subexpressos. Pelo menos um câncer hereditário, a doença de von Hippel-Lindau, na qual as pessoas desenvolvem múltiplos tipos de tumores, é causado pela mutação de um gene que afeta a angiogênese.

No desenvolvimento de muitos cânceres, o tumor primário dá origem a células que se disseminam para sítios distantes, produzindo tumores secundários. Esse processo de metástase é a causa da morte em 90% dos casos de câncer em humanos: é influenciado por mudanças celulares induzidas pela mutação somática. Como foi discutido na introdução deste capítulo, o gene *palladin*, quando sofre mutação, contribui para a metástase dos tumores pancreáticos. Ao usar microarranjos para medir os níveis da expressão gênica, os pesquisadores identificaram outros genes que são transcritos em uma taxa muito maior nas células metastáticas em comparação com as células não metastáticas. Por exemplo, um estudo detectou um conjunto de 95 genes que foram superexpressos ou subexpressos em uma população de células de câncer de mama que tinham metástases expressivas no pulmão, em comparação com uma população de células com uma população de células que tinham metástases discretas no pulmão. Os genes que contribuem para a metástase codificam componentes da matriz extracelular e do citoesqueleto. Outros codificam proteínas de adesão, que ajudam a manter as células unidas.

Avanços na tecnologia de sequenciamento possibilitam o sequenciamento completo do DNA de células tumorais para ver as diferenças de seus genomas em relação às células normais. Em um experimento, os pesquisadores sequenciaram o genoma inteiro de células de um tumor de câncer de mama com metástase e o compararam com o genoma de células não câncer da mesma pessoa. Eles também compararam o genoma do tumor metastático com o genoma do tumor primário (de onde originou a metástase), que foi removido da paciente 9 anos antes. Os pesquisadores encontraram 32 mutações somáticas diferentes nas regiões codificadoras dos genes das células tumorais, 19 das quais não foram detectadas no tumor primário. Esse achado sugere que o tumor metastático sofreu consideráveis mudanças genéticas na sua evolução de 9 anos a partir do tumor primário. Em contrapartida, outro estudo de metástase de câncer de mama encontrou apenas duas mutações que não estavam presentes no tumor primário, mas, nesse caso, a metástase tinha evoluído em apenas 1 ano.

Conceitos

As mutações nos genes codificadores dos componentes dos sistemas de reparo do DNA estão associadas com o câncer; essas mutações aumentam a taxa na qual as mutações são retidas e resultam em um número maior de mutações em proto-oncogenes, genes supressores de tumor e outros genes que contribuem para a proliferação celular. As mutações que permitem que a telomerase seja expressa nas células somáticas e as que afetam a vascularização e metástase também podem contribuir para o progresso do câncer.

Checagem dos conceitos 5

Qual tipo de mutação na telomerase está associado com as células do câncer?

a. Mutações que produzem uma forma inativa de telomerase.
b. Mutações que reduzem a expressão da telomerase.
c. Mutações que aumentam a expressão da telomerase.
d. Todas as opções anteriores.

MicroRNAs e câncer

Os microRNAs (miRNAs) são uma classe de pequenas moléculas de RNA que pareiam com sequências complementares no mRNA e degradam o mRNA ou inibem sua tradução (ver *Capítulo 14*). Considerando que os miRNAs são importantes para controlar a expressão gênica e o desenvolvimento, não é surpresa que eles também estejam associados com o desenvolvimento

do tumor. Muitas células tumorais exibem redução disseminada na expressão de muitos miRNAs. Os pesquisadores desenvolveram por engenharia genética células de tumor de camundongo que não tinham o maquinário para gerar miRNAs e descobriram que essas células mostravam um progresso acentuado do tumor quando implantadas nos camundongos. E interessante, esse efeito era observado apenas nas células que já tinham iniciado o desenvolvimento do tumor, sugerindo que os miRNAs são importantes nos estágios posteriores do progresso do tumor.

Níveis reduzidos de miRNAs podem contribuir para o câncer ao permitir que os oncogenes, que são normalmente controlados por miRNAs, sejam expressos em níveis elevados. Por exemplo, let-7 miRNA normalmente controla a expressão do oncogene *ras*, ao se ligar a sequências complementares da região 3′ não traduzida do mRNA e inibindo a tradução. Nas células do câncer de pulmão, os níveis de let-7 miRNA estão baixos, permitindo que a proteína Ras seja muito expressa, o que leva ao desenvolvimento do câncer de pulmão.

Um fator de transcrição chamado c-MYC é expresso com frequência em níveis elevados nas células do câncer. Evidências sugerem que c-MYC ajuda a direcionar a proliferação celular e o desenvolvimento do câncer. Entre outros efeitos, c-MYC se liga aos promotores dos genes do miRNA e reduz sua transcrição, reduzindo a abundância dos miRNAs. Alguns desses miRNAs são conhecidos por suprimir o desenvolvimento do tumor. A pesquisa mostra que, se, por manipulação genética, os miRNAs são expressos em níveis elevados, o desenvolvimento do tumor diminuiu. Todos esses achados sugerem que a expressão alterada dos miRNAs é importante para o câncer.

Vários miRNAs estão envolvidos no processo de metástase. Um miRNA particular, chamado miR-10b, está associado com a formação de tumores de mama metastáticos. Em um experimento, os investigadores manipularam uma linhagem de células de câncer de mama de modo que o miR-10b fosse superexpresso. Quando as células manipuladas foram injetadas nos camundongos, muitos deles desenvolveram tumores metastáticos. Ao contrário dos exemplos anteriores nos quais a expressão menor de miRNA está associada com o câncer, aqui os níveis elevados de miRNA promovem a disseminação das células do câncer. Outro estudo revelou que, nos humanos, os níveis de miR-10b eram elevados nos tumores metastáticos em comparação com os tumores nos pacientes livres de metástase. O miR-10b regula a expressão de vários outros genes, incluindo alguns conhecidos por suprimir a disseminação de células tumorais. Outros miRNAs inibem a metástase.

Projetos de genoma do câncer

Formado em 2008, o International Cancer Genome Consortium coordena os esforços para determinar as sequências genômicas dos tumores. Um dos objetivos do consórcio é sequenciar completamente 500 tumores a partir de cada um dos 50 tipos diferentes de câncer, junto com os genomas de tecidos normais das mesmas pessoas. Esse esforço está produzindo resultados importantes, revelando números e tipos de mutações associadas com determinados cânceres. A esperança é que sejam identificados novos genes responsáveis pelo câncer, o que levaria a uma melhor compreensão da natureza do câncer e sugeriria novos alvos para o tratamento do câncer. Outro projeto de pesquisa chamado The Cancer Genome Atlas (TCGA) começou nos EUA em 2005. Esse projeto visa obter uma análise genômica abrangente de cerca de 100 diferentes tipos de células de câncer, incluindo o sequenciamento de todos os éxons pelo genoma, assim como a caracterização da expressão do mRNA, metilação do DNA, variações no número de cópias e microRNAs das células tumorais.

Os genomas de vários tumores foram sequenciados e muito mais está sendo sequenciado. Por exemplo, o genoma inteiro de um carcinoma de células pequenas do pulmão (um tipo de câncer de pulmão) foi sequenciado em 2010 e comparado com o genoma de células normais da mesma pessoa. Foram identificadas mais de 22.000 mutações de pares de base no tumor, das quais 134 estavam em genes codificadores de proteínas. O tumor também tinha 58 rearranjos cromossômicos e 334 variações no número de cópias (ver *Capítulo 20*). Em outro estudo que foi parte do TCGA, os pesquisadores examinaram a expressão do mRNA, microRNAs, metilação do DNA e variações no número de cópias de 489 adenocarcinomas ovarianos e sequenciaram o DNA dos éxons de 316 dos tumores. Quase todos os tumores tinham mutações no *p53*, um gene supressor de tumor envolvido no reparo do DNA e controle do ciclo celular. As mutações no *BRCA1* e *BRCA2*, dois genes supressores de tumor, que também ocorrem no câncer de mama, ocorreram em 22% dos tumores. As mutações em sete outros genes ocorreram estatisticamente com maior frequência do que nas células normais. Os resultados sugeriram várias novas abordagens para o tratamento farmacológico do câncer de ovário.

Outra série de estudos sequenciou vários genes em amostras de gliomas malignos, uma forma incurável e mortal de câncer de cérebro. Os pesquisadores examinaram sequências de DNA, variações no número de cópias, metilação do DNA e expressão do RNA desses tumores. As análises revelaram mutações em vários genes que pareciam ser importantes no desenvolvimento dos tumores de glioma. Todos esses estudos genômicos estão fornecendo nova abordagem sobre a base genética do câncer.

O grande número de mutações encontradas em genomas de câncer pode ser dividido em dois tipos: **mutações condutoras (*driver*)** e **mutações passageiras (*passenger*)**. As mutações condutoras são as que impulsionam a evolução do câncer: elas contribuem diretamente para o desenvolvimento do câncer. As condutoras incluem mutações nos oncogenes, genes supressores de tumor, genes de reparo do DNA e outros tipos de genes de câncer discutidos neste capítulo. As passageiras são mutações que

surgem aleatoriamente no processo de desenvolvimento do tumor e não contribuem para o progresso do câncer. Muitas mutações passageiras estão em íntrons (regiões entre os genes) e outro DNA que não é transcrito e traduzido, mas elas também podem surgir dentro de genes que codificam proteínas. Um grande desafio é determinar quais das numerosas mutações encontradas nos tumores são condutoras e de fato contribuem para o desenvolvimento do câncer e quais são passageiras sem efeito.

23.3 As mudanças epigenéticas estão frequentemente associadas com o câncer

Mudanças epigenéticas, alterações na estrutura da cromatina que afetam a expressão gênica (ver *Capítulo 21*), são observadas em muitas células de câncer. Duas linhas amplas de evidência sugerem que as mudanças epigenéticas são importantes no progresso do câncer. Primeiro, os genes codificadores de proteínas que são importantes reguladoras das mudanças epigenéticas sofrem com frequência mutação em alguns cânceres. Por exemplo, quase 90% dos casos de linfoma folicular exibe mutações no gene *MLL2*, que codifica a enzima histona metiltransferase, essa enzima adiciona grupos metila ao DNA, um tipo de modificação epigenética que altera a estrutura da cromatina e afeta a transcrição. De forma semelhante, o gene *UTX*, que codifica uma histona desmetilase (enzima que remove os grupos metila das proteínas histona) sofre mutação em vários tipos diferentes de câncer.

Uma segunda linha de evidência, que sugere que as alterações epigenéticas são importantes no câncer, surge a partir de estudos genômicos recentes que comparam a estrutura da cromatina de células do câncer e células normais do mesmo indivíduo. Esses estudos descobriram que as células de câncer têm alterações importantes na metilação do DNA e na estrutura do DNA. Um tipo de alteração epigenética observada nas células de câncer é um nível geral de menos metilação de DNA (hipometilação). Como foi discutido no Capítulo 17, a metilação do DNA está associada com a repressão da transcrição. É possível que a hipometilação leve a transcrição dos oncogenes, que estimulam o câncer. Alguns indícios também sugerem que a hipometilação causa instabilidade cromossômica, uma marca de muitos tumores. As células tumorais de camundongos que foram geneticamente modificadas para ter a metilação do DNA reduzida mostram ganhos e perdas maiores de cromossomos, mas não está claro como a hipometilação pode causar instabilidade cromossômica.

Vários estudos observaram que, embora o nível geral de metilação do DNA seja menor nas células de câncer, algumas ilhas CpG específicas (ver *Capítulo 17*) têm metilação extra (são hipermetiladas). Por exemplo, um estudo descobriu que 5 a 10% de ilhas CpG normalmente não metiladas localizadas em promotores se tornam anormalmente metiladas nas células de câncer. Esse excesso de metilação pode inibir a transcrição de genes supressores de tumor, assim, estimulando o desenvolvimento do câncer. A metilação do promotor do gene *Apaf-1* é observada em qualquer célula maligna do melanoma. O *Apaf-1* produz apoptose das células com DNA danificado; a metilação de seu promotor reduz a expressão de *Apaf-1*, interrompendo o processo de apoptose e permitindo que as células anormais do câncer sobrevivam.

A pesquisa também demonstrou que as proteínas histona nos nucleossomos, a unidade fundamental da cromatina, estão modificadas de forma anormal nas células de câncer. A modificação das proteínas histona, incluindo metilação e acetilação, altera a estrutura da cromatina e afeta a transcrição. Os padrões globais de acetilação da histona estão alterados nas células de câncer. Entretanto, a acetilação não afeta apenas as histonas, mas também várias outras proteínas que podem estimular ou suprimir a divisão celular, então não está claro se o efeito da acetilação no câncer ocorre por mudanças na estrutura da cromatina. Os processos epigenéticos estão recebendo cada vez mais atenção dos pesquisadores de câncer porque podem ser sensíveis à terapia medicamentosa.

Conceitos

As mudanças epigenéticas, incluindo a metilação do DNA e modificação da histona, estão frequentemente associadas com o câncer.

Checagem dos conceitos 6

Acredita-se que a hipermetilação contribua para o câncer ao:

a. Inibir a replicação do DNA.
b. Inibir a expressão dos genes supressores de tumor.
c. Estimular a tradução dos oncogenes.
d. Estimular a telomerase.

23.4 O câncer colorretal surge em decorrência de mutação sequencial de vários genes

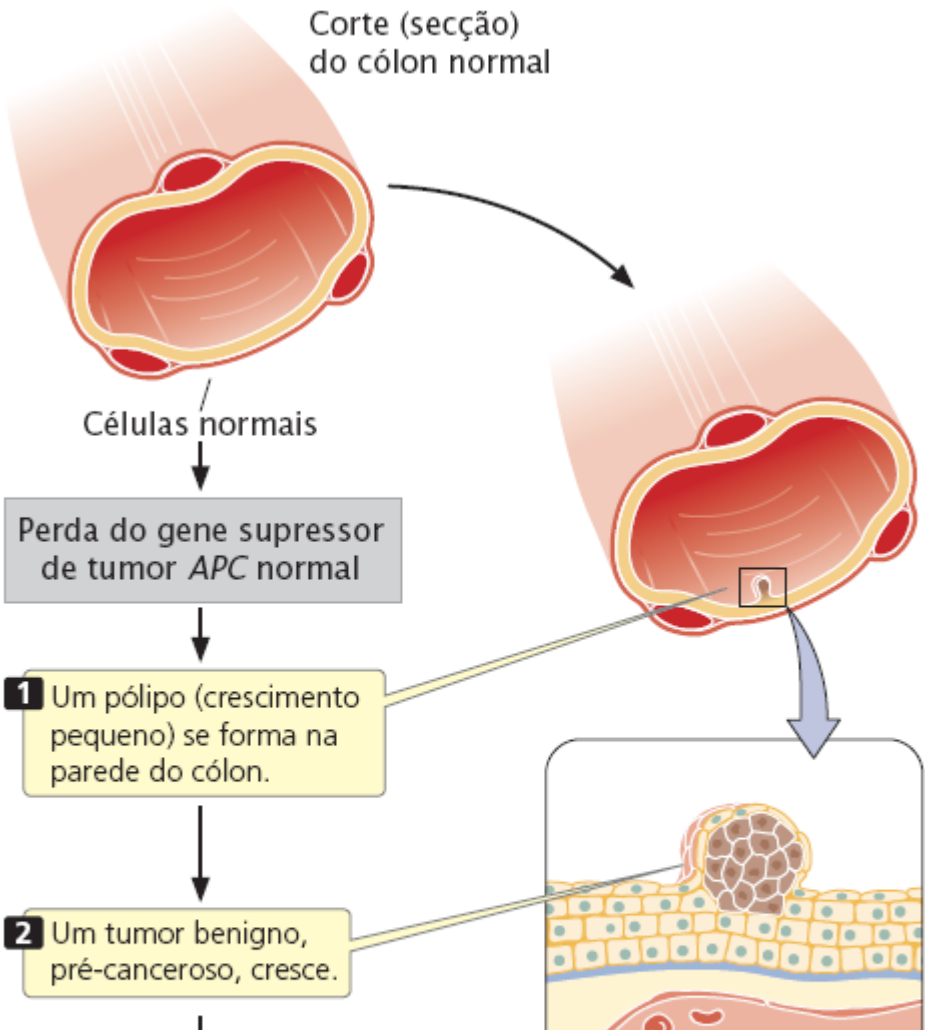
O câncer colorretal é um excelente exemplo de como o câncer surge pelo acúmulo de sucessivos defeitos genéticos. As mutações que contribuem para o câncer colorretal foram exaustivamente estudadas.

Esse câncer surge nas células que revestem o cólon e o reto. São diagnosticados mais de 143.000 novos casos de câncer colorretal nos EUA todo ano, onde esse câncer é responsável por mais de 51.000 mortes anualmente. Se detectado precocemente, o câncer colorretal pode ser tratado com sucesso; por esse motivo, existe muito interesse em identificar os eventos moleculares responsáveis pelos estágios iniciais desse câncer.

Acredita-se que o câncer colorretal se origine como tumores benignos chamados de pólipos adenomatosos (**Figura 23.10**). Inicialmente, esses pólipos são microscópicos, mas com o tempo, eles crescem e as células do pólipo adquirem características anormais de células cancerosas. Nos estágios posteriores da doença (avançados), o tumor pode invadir a camada de músculo que circunda o intestino e sofrer metástase. A evolução da doença é lenta, podendo levar de 10 a 35 anos para que um tumor benigno se desenvolva em um tumor maligno.

A maioria dos casos de câncer colorretal é esporádica, se desenvolve em pessoas sem história familiar da doença, mas algumas famílias apresentam uma clara predisposição genética para essa doença. Em uma forma de câncer de cólon hereditário, conhecido como polipose adenomatosa familiar, centenas ou milhares de pólipos se desenvolvem no cólon e no reto; se esses pólipos não forem removidos, um ou quase todos invariavelmente se tornam malignos.

Como os pólipos e os tumores do cólon e reto podem ser observados e removidos facilmente com um colonoscópio (um instrumento de fibra óptica usado para visualizar o interior do reto e do cólon), muito se sabe sobre o progresso do câncer colorretal e já foram identificados alguns dos genes responsáveis por sua evolução clonal. As mutações nesses genes são responsáveis pelas diferentes etapas do progresso desse câncer. Entre as etapas iniciais está uma mutação que inativa o gene *APC*, o que aumenta a taxa de divisão celular, levando a formação do pólipo (ver Figura 23.10). Uma pessoa com polipose adenomatosa familiar herda uma cópia defeituosa do gene *APC*, e os defeitos nesse gene estão associados com numerosos pólipos que surgem nos indivíduos com esse distúrbio. As mutações no *APC* também são encontradas nos pólipos que surgem em pessoas sem a polipose adenomatosa.



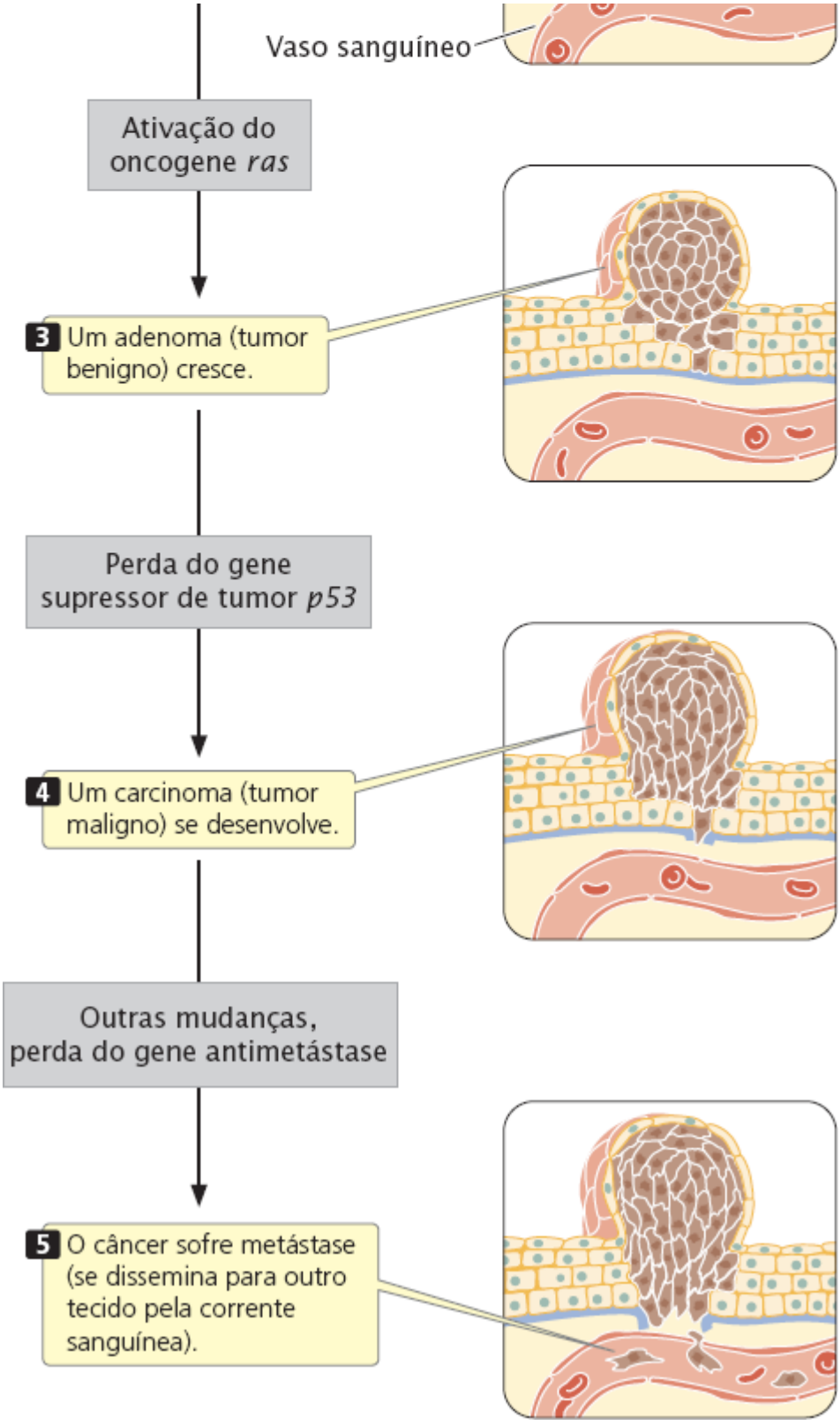


Figura 23.10 As mutações em múltiplos genes contribuem para o progresso do câncer colorretal.

As mutações no oncogene *ras* em geral ocorrem posteriormente, nos pólipos maiores compostos por células que adquiriram algumas mutações genéticas. Como foi discutido no início deste capítulo, o proto-oncogene *ras* normal é importante na via de transdução de sinal que restabelece os sinais a partir dos fatores de crescimento para o núcleo, onde o sinal estimula a divisão celular. Quando *ras* sofre mutação, a proteína que ele codifica continua a restabelecer um sinal estimulador para divisão celular mesmo quando o fator de crescimento está ausente.

Mutações em *p53* e outros genes aparecem posteriormente no progresso do tumor; essas mutações são raras em pólipos, mas comuns das células malignas. Cerca de 75% dos cânceres colorretais têm mutações no gene supressor de tumor *p53*. Como o *p53* evita a replicação de células com dano genético e controla a segregação cromossômica adequada, mutações nesse gene podem permitir que uma célula adquira rapidamente mais mutações em genes e cromossomos, o que contribui para mais proliferação e invasão nos tecidos circundantes.

A sequência de etapas destacada não é a única rota para o câncer colorretal e as mutações não precisam ocorrer na ordem apresentada aqui. Entretanto, essa sequência é uma via comum pela qual as células do cólon e reto se tornam cancerosas.

23.5 Mudanças no número e na estrutura do cromossomo estão frequentemente associadas com câncer

A maioria dos tumores tem células com mutações cromossômicas. Por muitos anos, os geneticistas se perguntavam se essas mutações cromossômicas eram a causa ou o resultado do câncer. Alguns tipos de tumores estão associados com mutações cromossômicas específicas; por exemplo, a maioria dos casos de leucemia mielocítica crônica está associada com uma translocação recíproca entre os cromossomos 22 e 9. Esses tipos de associações sugerem que as mutações cromossômicas contribuem para a causa do câncer. Até o momento, muitos cânceres não estão associados com tipos específicos de anomalias cromossômicas e agora sabemos que mutações *gênicas* individuais contribuem para muitos tipos de câncer. Contudo, a instabilidade cromossômica é uma característica geral das células cancerosas, fazendo com que elas acumulem mutações de cromossomos, que afetam os genes individuais e podem contribuir para o processo do câncer. Assim, parece que as mutações cromossômicas são tanto a *causa* quanto o *resultado* do câncer.

Pelo menos três tipos de rearranjos cromossômicos, deleções, inversões e translocações estão associados com determinados tipos de câncer. As deleções podem surgir na perda de um ou mais genes supressores de tumor. As inversões e translocações contribuem para o câncer de várias formas. Primeiro, os pontos de quebra cromossômica que acompanham essas mutações podem ficar dentro dos genes supressores de tumor, prejudicando sua função e levando à proliferação celular. Segundo, as translocações e inversões podem produzir sequências a partir de dois genes diferentes, gerando uma proteína de fusão que estimula algum aspecto do progresso do câncer.

As proteínas de fusão são observadas na maioria dos casos de leucemia mielocítica crônica, que afeta as células da medula óssea. A maioria dos pacientes com esse tipo de leucemia tem uma translocação recíproca entre o braço longo do cromossomo 22 e a ponta do braço longo do cromossomo 9 (**Figura 23.11**). Essa translocação produz um cromossomo 22 encurtado, chamado de cromossomo Philadelphia; e recebeu este nome porque foi descoberto pela primeira vez na Filadélfia. O cromossomo 9 normal tem um gene chamado *c-ABL* na sua extremidade, que é potencial causador de câncer. Como resultado da translocação, parte do gene *c-ABL* é fundida com o gene *BRC* do cromossomo 22. A proteína produzida por esse gene da fusão *BCR-cABL* é muito mais ativa que a proteína produzida pelo gene *c-ABL* normal; essa proteína estimula a divisão celular aumentada e desregulada e, consequentemente, leva a leucemia.

Um terceiro mecanismo pelo qual os rearranjos cromossômicos podem produzir câncer é a transferência de um potencial gene causador de câncer para uma nova localização, onde é ativado por diferentes sequências reguladoras. O linfoma de Burkitt é um câncer de linfócitos B produtores de anticorpos. Muitas pessoas com o linfoma de Burkitt tem uma translocação recíproca entre o cromossomo 8 e o cromossomo 2, 14 ou 22 (**Figura 23.12**). Essa translocação realoca um gene chamado *c-MYC* da ponta do cromossomo 8 para uma posição no cromossomo 2, 14 ou 22 que está próxima de um gene que codifica uma proteína imunoglobulina. Nessa nova localização, o *c-MYC*, um gene causador de câncer, fica sob controle de sequências reguladoras que normalmente ativam a produção de imunoglobulinas e assim o *c-MYC* é expresso nos linfócitos B. A proteína *c-MYC* estimula a divisão dos linfócitos B e leva ao linfoma de Burkitt.

Conceitos

Muitos tumores têm vários tipos de mutações cromossômicas. Alguns tumores estão associados com deleções, inversões e translocações específicas. As deleções podem eliminar ou inativar genes que controlam o ciclo celular, as inversões e translocações podem causar quebras em genes que suprimem tumores, fundir genes para produzir proteínas causadoras de câncer ou deslocar genes para novos locais, onde eles estão sob a influência de diferentes sequências reguladoras.

Checagem dos conceitos 7

A leucemia mielocítica crônica está associada com qual tipo de rearranjo cromossômico?

a. Duplicação.
b. Deleção.
c. Inversão.
d. Translocação.

A maioria dos tumores avançados tem células que exibem uma variedade incrível de anomalias cromossômicas, incluindo cromossomos extras, cromossomos ausentes e rearranjos cromossômicos (**Figura 23.13**). Alguns pesquisadores de câncer acreditam que o câncer é iniciado quando ocorrem mudanças genéticas que fazem o genoma ficar instável, gerando numerosas anomalias cromossômicas que então alteram a expressão dos oncogenes e os genes supressores de tumor.

Já foram identificados vários genes que contribuem para a instabilidade genômica e levam a perda ou cromossomos extras (aneuploidia). A aneuploidia nas células somáticas em geral surge quando os cromossomos não se separam adequadamente na mitose. As células normais têm um ponto de verificação da montagem do fuso que monitora a montagem adequada do fuso mitótico (ver p. 603); se os cromossomos não estão adequadamente fixos aos microtúbulos na metáfase, o início da anáfase é bloqueado. Algumas células aneuploides de câncer têm alelos mutantes para os genes que codificam proteínas que participam do ponto de verificação; nessas células, a anáfase é iniciada apesar da montagem incorreta do fuso ou da ausência desse e surgem anomalias cromossômicas. Por exemplo, as mutações em *RB* aumentam a aneuploidia ao aumentar a expressão de uma proteína chamada Mad2, que é um componente crítico do ponto de verificação da montagem do fuso.

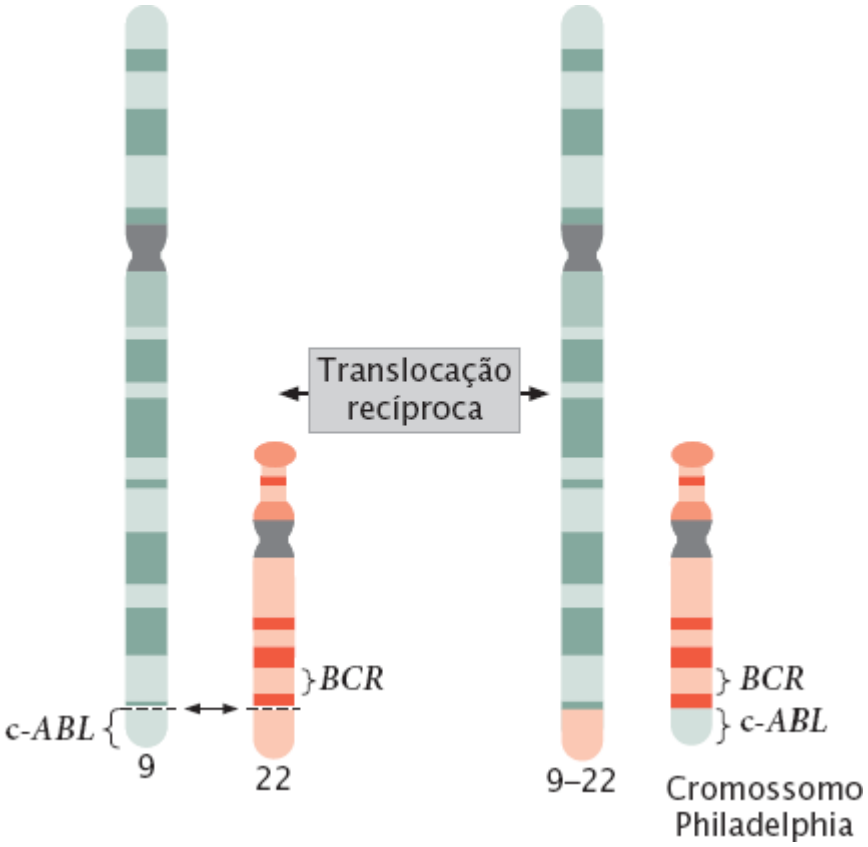


Figura 23.11 Uma translocação recíproca entre os cromossomos 9 e 22 é responsável pela leucemia mielocítica crônica.

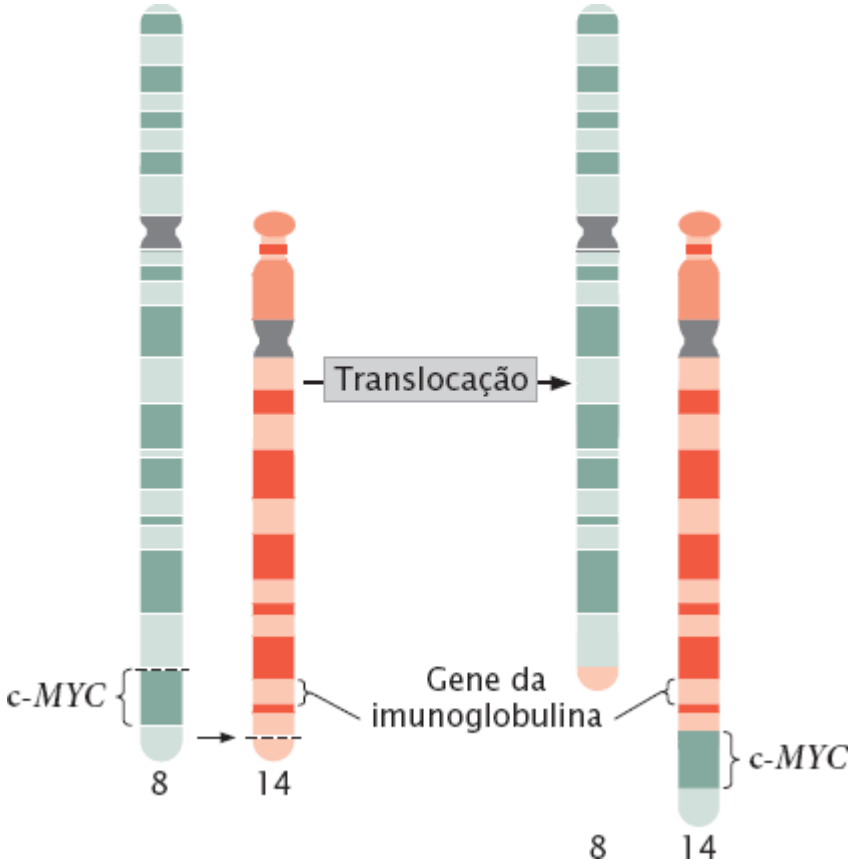


Figura 23.12 Uma translocação recíproca entre os cromossomos 8 e 14 é responsável pelo linfoma de Burkitt.

As mutações em genes codificadores de partes do aparato do fuso também podem contribuir para a segregação anormal e levar a anomalias cromossômicas. O *APC* é um gene supressor de tumor que sofre mutação em células do câncer do cólon. Esse gene tem várias funções, uma das quais é interagir com as extremidades dos microtúbulos que se associam com o cinetócoro. As células de camundongo em divisão com cópias defeituosas do gene *APC* dão origem a células com muitos defeitos cromossômicos.

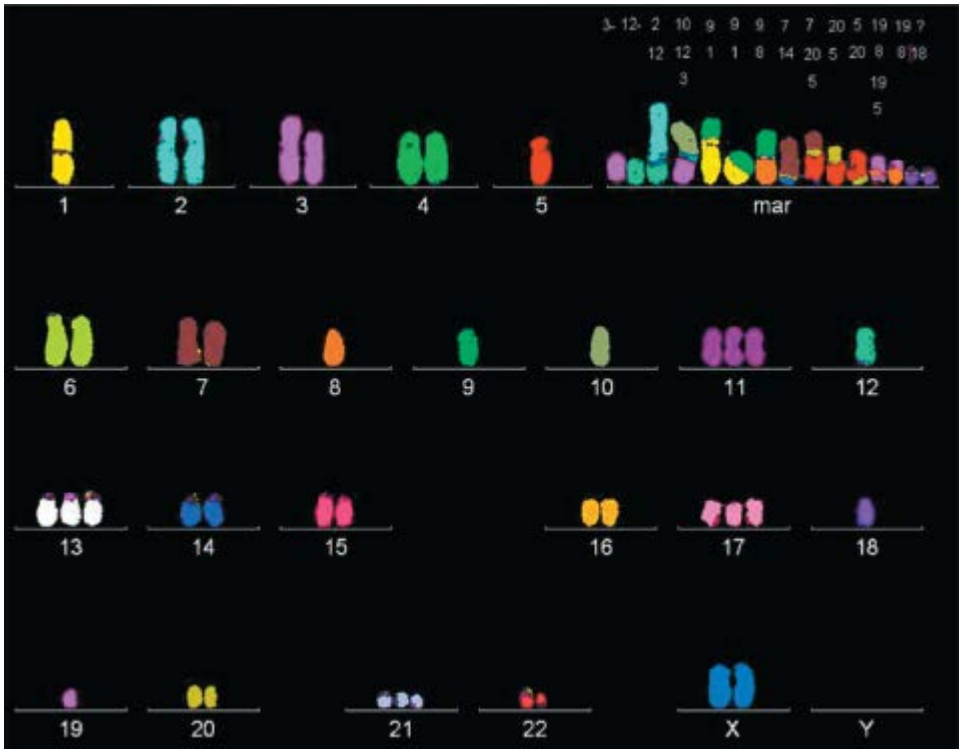


Figura 23.13 As células de câncer têm anomalias cromossômicas, incluindo cromossomos extras, cromossomos ausentes e rearranjos cromossômicos. Aqui estão apresentados os cromossomos de uma célula de câncer de cólon, com numerosas anomalias cromossômicas. Por comparação, observe um cariótipo normal na Figura 2.5 (Cortesia de Dr. Peter Duesberg, UC Berkeley.)

O gene supressor de tumor *p53*, além de controlar a apoptose, participa da duplicação do centrossomo, necessário para a formação adequada do fuso e para a segregação do cromossomo. Normalmente, o centrossomo duplica uma vez a cada ciclo celular. Se o gene *p53* sofrer mutação, entretanto, o centrossomo pode sofrer duplicações extras, resultando na segregação desigual dos cromossomos. Dessa forma, a mutação do gene *p53* pode gerar mutações cromossômicas que contribuem com o câncer. O gene *p53* também é um gene supressor de tumor que impede a divisão celular quando o DNA está danificado.

Resolva o Problema 32

23.6 Os vírus estão associados com alguns cânceres

Como já foi mencionado neste capítulo, os vírus são responsáveis por vários cânceres em animais, e existem indícios de que alguns vírus contribuem para alguns cânceres em humanos (**Quadro 23.6**). Por exemplo, cerca de 95% das mulheres com câncer do colo uterino estão infectadas por **papilomavírus humano** (HPV). Da mesma forma, a infecção com o vírus responsável pela hepatite B aumenta o risco de câncer hepático em algumas pessoas. O vírus Epstein-Barr, responsável pela mononucleose, está ligado a vários tipos de câncer que são prevalentes em parte da África, incluindo o linfoma de Burkitt.

Retrovírus e câncer. Muitos dos vírus que causam câncer em animais são retrovírus; já observamos como os estudos do retrovírus do sarcoma de Rous em galinhas levaram à identificação dos oncogenes nos humanos. Os retrovírus podem causar câncer por mutação e rearranjo dos genes do hospedeiro, convertendo os proto-oncogenes em oncogenes (**Figura 23.14 A**). Outra forma na qual os vírus podem contribuir para o câncer é alterar a expressão dos genes do hospedeiro (**Figura 23.14 B**). Os retrovírus têm fortes promotores que garantem que seu próprio material genético seja transcrito pela célula hospedeira. Se o pró-vírus se inserir próximo a um proto-oncogene, os promotores virais podem estimular níveis elevados de expressão do proto-oncogene, levando à proliferação celular.

Quadro 23.6 Alguns cânceres em humanos associados com vírus.

Vírus	Câncer
Papilomavírus humano (HPV)	Cânceres do colo uterino, pênis e vulvar
Vírus da hepatite B	Câncer de fígado
Vírus 1 da leucemia de células T humana (HTLV-1)	Leucemia da célula T em adultos
Vírus 2 da leucemia de células T humana (HTLV-2)	Tricoleucemia
Vírus Epstein-Barr	Linfoma de Burkitt, câncer nasofaríngeo, linfoma de Hodgkin
Herpes-vírus humano	Sarcoma de Kaposi
Poliomavírus das células de Merkel	Carcinoma das células de Merkel

Nota: Algumas dessas associações entre câncer e vírus existem apenas em algumas populações e áreas geográficas.

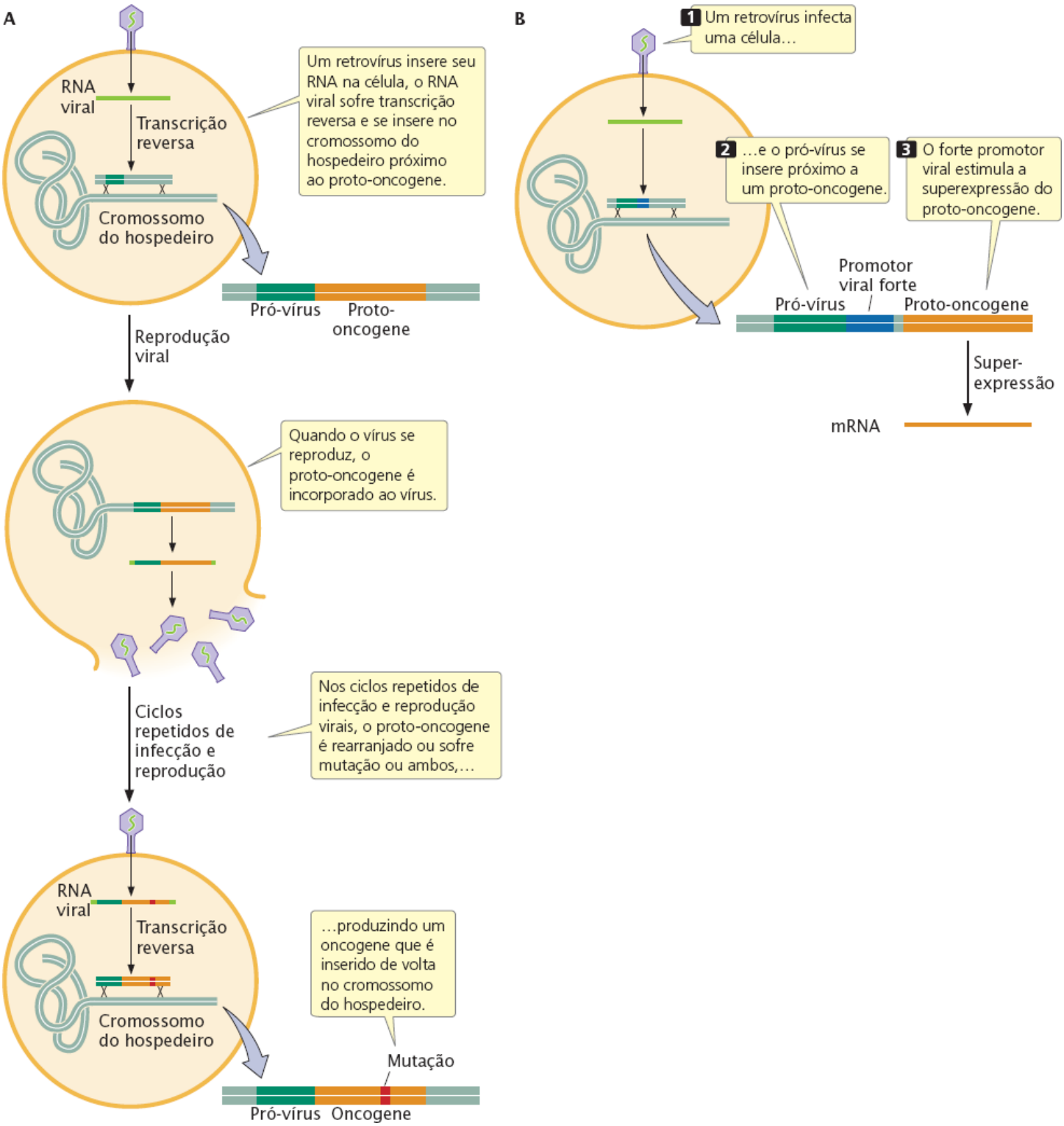


Figura 23.14 Os retrovírus causam câncer ao (A) promover mutação e rearranjo de proto-oncogenes ou (B) inserir promotores fortes próximo dos proto-oncogenes.

Existem apenas alguns retrovírus que causam câncer nos humanos. O HTLV-1, o primeiro retrovírus humano descoberto, está associado com a leucemia de células T em adultos. Outros cânceres humanos estão associados com vírus DNA, que, como os retrovírus, se integram ao cromossomo hospedeiro, mas diferente dos retrovírus, não usam a transcrição reversa. Por exemplo, o HPV é um vírus DNA que está fortemente associado com o câncer do colo uterino.

Papilomavírus humano e câncer do colo uterino. O papilomavírus humano produz verrugas e outros tipos de tumores benignos das células epiteliais. São conhecidos mais de 100 diferentes tipos de HPV; cerca de 30 destes são sexualmente transmitidos e causam verrugas genitais. Alguns deles estão associados com o câncer do colo uterino. Nos EUA, 70% dos casos desse câncer são causados por HPV-16 e HPV-18. Esses vírus causam o câncer do colo uterino ao produzir proteínas que se fixam e inativam *RB* e *p53*, duas proteínas com papeis críticos na regulação do ciclo celular. Quando essas proteínas são inativadas, as células são estimuladas a progredir no ciclo celular e se dividir sem os controles normais que evitam a proliferação celular.

Cerca de 75% das mulheres sexualmente ativas nos EUA estão infectadas pelo HPV, mas apenas um pequeno número desenvolverá esse tipo de câncer. O risco de infecção pelo HPV pode ser reduzido ao limitar o número de parceiros sexuais e com uma vacina aprovada pela U.S. Food and Drug Administration em 2006. Essa vacina é muito eficiente contra quatro dos HPVs mais comuns associados com o câncer do colo uterino. Apesar do grande número de mulheres infectadas com HPVs, a incidência desse câncer nos EUA caiu 75% nos últimos 40 anos. O principal motivo para esse declínio é o uso difundido do esfregaço de Papanicolaou, que detecta os estágios iniciais do câncer, assim como a displasia cervical, um crescimento pré-canceroso que pode ser removido antes que se torne um câncer.

Embora seja raro nos EUA, o câncer de colo uterino é a segunda principal causa de câncer em mulheres pelo mundo, com altas incidências em muitos países em desenvolvimento, como os da África subsaariana, sul da Ásia e Américas do Sul e Central. Existe uma estimativa de que 510.000 mulheres no mundo desenvolvam este câncer todos os anos e 288.000 destas pereçam. O motivo primário para as altas incidência e taxa de morte nas regiões em desenvolvimento é a falta de acesso aos procedimentos de triagem do câncer como o esfregaço de Papanicolaou. A disponibilidade da vacina contra esse câncer nos países em desenvolvimento poderia reduzir muito a incidência da doença.

Conceitos

Os vírus contribuem para alguns cânceres nos humanos por meio de mutação e rearranjo dos genes do hospedeiro que então contribuem para a proliferação celular ou pela alteração da expressão dos genes do hospedeiro.

Resumo dos conceitos

- O câncer é fundamentalmente um distúrbio genético, surgindo a partir de mutações somáticas em múltiplos genes que afetam a divisão e a replicação celular. Se uma ou mais mutações são herdadas, são necessárias poucas mutações adicionais para surgir o câncer
- Uma mutação que permite uma célula se dividir rapidamente fornece a esta célula uma vantagem do crescimento; essa célula dá origem a um clone de células com a mesma mutação. Nesse clone, ocorrem outras mutações que fornecem vantagens adicionais de crescimento e as células com essas mutações adicionais se tornam dominantes no clone. Dessa forma, o clone evolui
- Os fatores ambientais são importantes no desenvolvimento de muitos cânceres ao aumentar a taxa das mutações somáticas
- Os oncogenes são cópias mutantes dominantes dos genes normais (proto-oncogenes) que estimulam a divisão celular. Normalmente, os genes supressores de tumor inibem a divisão celular; as mutações recessivas nesses genes podem contribuir para o câncer. Em alguns casos, a mutação de um único alelo de um gene supressor de tumor é suficiente para causar o câncer, um fenômeno chamado de haploinsuficiência
- O ciclo celular é controlado por ciclinas e quinases dependentes de ciclina. As mutações nos genes que controlam o ciclo celular estão associadas com o câncer
- As vias de transdução de sinal conduzem sinais externos para respostas intracelulares. Essas vias são compostas por uma série de proteínas que são ativadas e inativadas em uma cascata de reações. As mutações nos componentes das vias de transdução de sinal atrapalham o ciclo celular e podem contribuir para o câncer
- Defeitos nos genes de reparo do DNA aumentam a taxa geral de mutação de outros genes, levando a defeitos nos proto-oncogenes e genes supressores de tumor que podem contribuir com o progresso do câncer
- As mutações nas sequências que regulam a telomerase permitem que as células se dividam indefinidamente, contribuindo com o progresso do câncer
- O progresso do câncer também é afetado por mutações nos genes que promovem a vascularização e a disseminação dos tumores
- Muitas células tumorais exibem uma redução disseminada na expressão de muitos miRNAs, sugerindo que uma redução no controle do miRNA na expressão gênica pode ser importante no progresso do tumor
- As mudanças epigenéticas na estrutura da cromatina, incluindo metilação do DNA e modificação da histona, estão associadas com o câncer
- O câncer colorretal é um sistema modelo para compreender o progresso do tumor nos humanos. As mutações iniciais estimulam a divisão celular, gerando um pequeno pólipo benigno. Mutações adicionais permitem que o pólipo cresça, invada a camada muscular do intestino e se dissemine para outros tecidos. As mutações em genes específicos afetam diferentes estágios desse progresso
- Alguns tumores estão associados com mutações cromossômicas específicas, incluindo deleções, inversões e translocações

cromossômicas. As deleções podem causar o câncer ao remover ou romper os genes que suprimem os tumores, as inversões e translocações podem romper esses genes ou eles podem mover os genes para posições próximas a diferentes sequências reguladoras, que alteram sua expressão

- As mutações em alguns genes causam ou permitem que os cromossomos se separem de forma inadequada, levando a aneuploidia que pode contribuir com o câncer
- Os vírus estão associados com alguns cânceres; eles contribuem para a proliferação celular ao causar mutação e rearranjo dos genes do hospedeiro ou ao alterar a expressão desses genes.

Termos importantes

Apoptose
Ciclina
Evolução clonal
Gene supressor de tumor
Haploinsuficiência
Metástase
Mutação condutora (*driver*)
Mutação passageira (*passenger*)
Oncogene
Papilomavírus humano (HPV)
Perda da heterozigosidase
Proto-oncogene
Quinase dependente de ciclina (CDK)
Tumor maligno
Via de transdução de sinal

Respostas da Checagem dos conceitos

1. O retinoblastoma resulta de pelo menos dois defeitos genéticos separados, ambos necessários para o câncer se desenvolver. Nos casos esporádicos, duas mutações sucessivas em uma única célula têm de ocorrer, o que é improvável e, portanto, afeta apenas um olho. Nas pessoas que herdaram uma das duas mutações necessárias, cada célula tem essa mutação, e então uma única mutação adicional é tudo que se precisa para o câncer se desenvolver. Considerando os milhões de células em cada olho, existe uma elevada probabilidade de que ocorrerá uma segunda mutação em pelo menos uma célula de cada olho, produzindo tumores bilaterais e a herança desse tipo de retinoblastoma.
2. Os oncogenes têm um efeito estimulador sobre a proliferação celular. As mutações nos oncogenes são dominantes porque uma mutação em uma única cópia de gene é suficiente para produzir um efeito estimulante. Os genes supressores de tumor inibem a proliferação celular. Os genes supressores de tumor mutantes em geral são recessivos, porque ambas as cópias têm de ser mutantes para remover toda a inibição.
3. d.
4. a.
5. c.
6. b.
7. d.

Problema desenvolvido

Problema

Em algumas células do câncer, um gene específico foi duplicado várias vezes. É provável que esse gene seja um oncogene ou um gene supressor de tumor? Apresente sua justificativa.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
Se é provável que o gene seja um oncogene ou um gene supressor de tumor e por quê.

Quais informações são fornecidas para solucionar o problema?

Nas células do câncer, esse gene foi ampliado muitas vezes.

Para a solução deste problema, revise:
Oncogenes e genes supressores de tumor, na Seção 23.2.

Etapas para a solução

É provável que o gene seja um oncogene. Os oncogenes estimulam a proliferação celular e atuam de forma dominante. Portanto, cópias extras de um oncogene levarão a proliferação celular e câncer. Os genes supressores de tumor, por outro lado, suprimem a proliferação celular e atuam de forma recessiva, uma única cópia desse gene é suficiente para evitar a proliferação celular. Portanto, cópias extras do gene supressor do tumor não produziriam o câncer.

Lembrete: Um oncogene é um acelerador da divisão celular, enquanto um gene supressor de tumor tem ação oposta.

Questões de compreensão

Seção 23.1

1. Quais tipos de evidências indicam que o câncer surge a partir de mudanças genéticas?
2. Como o câncer é diferente da maioria dos outros tipos de doenças genéticas?
3. Descreva a hipótese de dois eventos de Knudson para o retinoblastoma e descreva como ela ajuda a explicar os casos de retinoblastoma unilateral e bilateral.
4. Explique resumidamente como o câncer surge a partir da evolução clonal.

Seção 23.2

5. Qual é a diferença entre um oncogene e um gene supressor de tumor? Forneça alguns exemplos das funções dos proto-oncogenes e supressores de tumor nas células normais.
6. O que é haploinsuficiência? Como ela pode afetar o risco de ter câncer?
7. Qual a diferença entre as ciclinas e CDKs? Como elas interagem para controlar o ciclo celular?
8. Descreva resumidamente os eventos que controlam o progresso das células pelo ponto de verificação G_1/S no ciclo celular.
9. Descreva resumidamente os eventos que controlam o progresso das células pelo ponto de verificação G_2/M no ciclo celular.
10. O que é uma via de transdução de sinal? Por que as mutações nos componentes nas vias de transdução de sinal estão frequentemente associadas com o câncer?
11. Como a proteína Ras é ativada e inativada?
12. Por que as mutações nos genes que codificam as enzimas de reparo do DNA produzem predisposição ao câncer?
13. Qual o papel dos telômeros e da telomerase no progresso do câncer?

Seção 23.3

14. Qual é a diferença entre uma mudança epigenética e uma mutação?
15. Qual a relação da metilação do DNA com o câncer?

Seção 23.4

16. Descreva resumidamente algumas das mudanças genéticas comumente associadas com o progresso do câncer colorretal.

Seção 23.5

*17. Explique como as deleções, inversões e translocações cromossômicas podem causar câncer.
18. Descreva resumidamente como o cromossomo Philadelphia causa a leucemia mielocítica crônica.
19. O que é instabilidade genômica? Forneça algumas formas pelas quais a instabilidade genômica pode surgir.

Seção 23.6

*20. Como os vírus contribuem para o câncer?

Questões e problemas aplicados

Introdução

21. Quais características do heredograma apresentado na Figura 23.1 sugerem que o câncer pancreático nessa família é herdado como um traço autossômico dominante?

Seção 23.1

22. Se o câncer é fundamentalmente uma doença genética, como um fator ambiental como o tabagismo pode causar câncer?

*23. Tanto genes como fatores ambientais contribuem para o câncer. O Quadro 23.2 mostra que o câncer de próstata é 30 vezes mais comum nos caucasianos de Utah do que os chineses de Xangai. Descreva resumidamente como você poderia determinar se essas diferenças de incidência de câncer de próstata se devem a diferenças na composição genética das duas populações ou a diferenças em seus ambientes.

*24. Um casal tem um filho com retinoblastoma bilateral. A mãe não tem câncer, mas o pai tem retinoblastoma unilateral e seu irmão tem retinoblastoma bilateral.
   a. Se esse casal tiver outra criança, qual é a probabilidade de que essa criança terá retinoblastoma?
   b. Se a próxima criança tiver retinoblastoma, ele será unilateral ou bilateral?
   c. Explique por que o retinoblastoma do pai é unilateral enquanto no filho e no irmão é bilateral.

Seção 23.2

*25. O gene *palladin*, envolvido no câncer pancreático (ver a introdução deste capítulo), é considerado um oncogene. Quais das suas características sugerem que ele é um oncogene em vez de um gene supressor de tumor?

26. As mutações no gene *RB* estão frequentemente associadas com câncer. Explique como uma mutação que resulta em uma proteína RB não funcional contribui para o câncer.

27. As células em um tumor têm cópias mutantes de um gene específico que promove o crescimento tumoral. A terapia gênica pode ser usada para introduzir uma cópia normal desse gene nas células do tumor. Você esperaria que essa terapia fosse eficiente se o gene mutante for um oncogene? Um gene supressor de tumor? Explique sua resposta.

28. Qual seria o efeito sobre o ciclo celular de um fármaco que inibe cada um dos componentes a seguir?
   a. MPF.
   b. ciclina-E-CDK.
   c. ciclina-D-CDK.

29. Qual seria o efeito de um fármaco que inibe a degradação da ciclina B?

Seção 23.3

ANÁLISE DE DADOS

30. David Seligson e colaboradores examinaram os níveis da modificação da proteína histona em tumores de próstata e sua associação com os desfechos clínicos (D. B. Seligson *et al*. 2005. *Nature* 435:1262-1266). Eles usaram anticorpos para marcar a acetilação em três sítios diferentes e corar a metilação em dois sítios diferentes nas proteínas histona. Eles descobriram que o grau de acetilação e metilação da histona ajudava a prever se o câncer de próstata retornaria dentro de 10 anos nos pacientes que tiveram o tumor removido. Explique como a acetilação e a metilação podem estar associadas com a recorrência do tumor no câncer de próstata. (Dica: ver *Capítulo 17*.)

31. Alguns cânceres são tratados com fármacos que desmetilam o DNA. Explique como esses fármacos funcionam. Você acredita que os genes causadores de câncer que respondam à desmetilação sejam oncogenes ou genes supressores de tumor? Explique sua justificativa.

Seção 23.5

*32. Alguns cânceres estão associados com a deleção de uma parte particular de um cromossomo. A região eliminada tem um oncogene ou um gene supressor de tumor? Explique.

Seção 23.6

33. Considere que o pró-vírus na Figura 23.14 se insira logo *upstream* de um gene supressor de tumor. Qual seria sua probabilidade de causar câncer? Por que sim ou por que não?

Seção 23.2

34. Muitas células de câncer são imortais (se dividem indefinidamente) porque têm mutações que permitem que a telomerase seja expressa. Como esse conhecimento poderia ser usado para desenvolver fármacos anticâncer?
35. A síndrome de Bloom é uma doença autossômica recessiva que exibe haploinsuficiência. Um levantamento recente mostrou que as pessoas heterozigotas para mutações no *locus BLM* correm maior risco de câncer de cólon. Imagine que você é um conselheiro genético. Você recebe o encaminhamento de uma mulher jovem cuja mãe tem síndrome de Bloom, mas o pai não tem história familiar da síndrome de Bloom. Ela pergunta se existe possibilidade de ter qualquer problema de saúde associado com sua história familiar da síndrome de Bloom. Que orientação você daria para essa pessoa?
36. Imagine que você descobre uma família numerosa na qual o câncer de bexiga é herdado como um traço autossômico dominante. Descreva resumidamente uma série de estudos que você poderia conduzir para identificar o gene que causa câncer de bexiga nessa família.

24

Genética Quantitativa

Óleo de milho e genética quantitativa

Em 2012, a população mundial ultrapassou 7 bilhões de pessoas. A Organização das Nações Unidas calcula que em 2050, ela terá aumentado mais 2 bilhões: alimentar estes bilhões de pessoas a mais será o principal desafio da agricultura nas próximas décadas. As plantações terão que fornecer mais calorias e nutrientes necessários para a futura população do mundo. Com a redução do fornecimento de petróleo e a preocupação com o aquecimento global, as plantas estão sendo cada vez mais usadas como fontes de biocombustíveis, aumentando ainda mais a demanda nas colheitas.

Para ajudar a satisfazer a demanda por culturas com maior rendimento, os agricultores estão usando as últimas técnicas genéticas na sua busca para desenvolver culturas com maior rendimento e mais eficientes. O poder desta abordagem é demonstrado pela pesquisa que procura aumentar o teor de óleo do milho. O teor de óleo no milho é herdado, mas a herança é mais complexa do que as características que estudamos até o momento; não é apenas uma única característica monogênica, como o formato da semente nas ervilhas. Numerosos genes e fatores ambientais contribuem para o teor de óleo no milho. Para características como o teor de óleo, vários *loci* interagem e sua expressão é afetada por fatores ambientais. Como a herança de uma característica complexa como o teor de óleo pode ser estudada? É possível prever o teor de óleo de uma planta com base na sua criação? A resposta é sim, pelo menos em parte, mas estas perguntas não podem ser discutidas com os métodos que usamos para as características genéticas simples. Por sua vez, devemos usar procedimentos estatísticos que foram desenvolvidos para analisar características complexas. A análise genética de características complexas como o teor de óleo do milho é conhecida como **genética quantitativa**.

Em 2008, os geneticistas usaram uma combinação de genética quantitativa e técnicas moleculares para identificar um gene-chave que controla o teor de óleo no milho. Primeiro, eles fizeram cruzamentos entre plantas com alto e baixo teor de óleo para identificar as regiões cromossômicas importantes para determinar a produção do óleo. As regiões cromossômicas com os genes que influenciam um traço quantitativo são chamadas ***loci* de traço quantitativo** (QTLs). Por meio destes cruzamentos, os geneticistas localizaram vários QTLs que afetam o teor de óleo; um deles estava no cromossomo 6 do milho.

O mapeamento genético de escala fina estreitou o QTL para uma pequena região de 4,2 centiMorgans no cromossomo 6. Os pesquisadores sequenciaram o DNA da região e descobriram que ele tinha cinco genes, um dos quais era *DGAT1-2*, conhecido por produzir uma enzima que catalisa a etapa final em uma via de biossíntese de triacilglicerol. O DNA do gene *DGAT1-2* em uma cepa de alta produção de óleo tinha uma inserção de um códon que adicionava fenilalanina à enzima, uma inserção que estava ausente na cepa de produção reduzida de óleo. Os pesquisadores confirmaram o efeito do códon adicional de fenilalanina na produção de óleo ao produzir milho transgênico com o códon extra; a produção de óleo destas cepas transgênicas aumentou em comparação com as cepas transgênicas sem o códon extra. E interessante, o códon extra de fenilalanina está presente em espécies selvagens de milho, sugerindo que o códon foi perdido durante o processo de domesticação ou subsequente criação de variedades modernas.

Esta pesquisa sugere que o teor de óleo de milho e de outras plantas pode ser aumentado por modificação genética de seus genes *DGAT1-2* para ter o códon extra da fenilalanina. Outros estudos que combinam de forma semelhante as análises quantitativa e molecular levaram à identificação de genes que aumentam o teor de vitamina A do arroz e aumentam a produção de açúcar nos tomates.

Este capítulo discute a análise genética de características complexas como o teor de óleo do milho. Vamos começar analisando as diferenças entre as características quantitativas e qualitativas e por que a expressão de algumas características varia de forma contínua. Vamos ver como as características quantitativas são influenciadas por muitos genes, cada um com um pequeno efeito no fenótipo. A seguir, vamos examinar os procedimentos estatísticos para descrever e analisar as características quantitativas. Vamos analisar o quanto a variação fenotípica pode ser atribuída a influências genéticas e ambientais. Finalmente, examinaremos os efeitos da seleção sobre as características quantitativas.

24.1 As características quantitativas variam continuamente e muitas são influenciadas por alelos em múltiplos *loci*

As características qualitativas, ou descontínuas, têm apenas alguns fenótipos distintos (**Figura 24.1 A**); estas características são os tipos estudados por Mendel (p. ex., ervilhas lisas e rugosas) e são o foco de nossa atenção até o momento. Entretanto, muitas características variam continuamente ao longo de uma escala de medida com muitos fenótipos sobrepostos (**Figura 24.1 B**). Elas são chamadas de *características contínuas*, e também são chamadas de *características quantitativas* porque o fenótipo de qualquer indivíduo tem de ser descrito pela medida quantitativa. Os exemplos de características quantitativas incluem altura, peso e pressão arterial nos humanos, taxa de crescimento nos camundongos, peso das sementes nas plantas e produção de leite no gado.

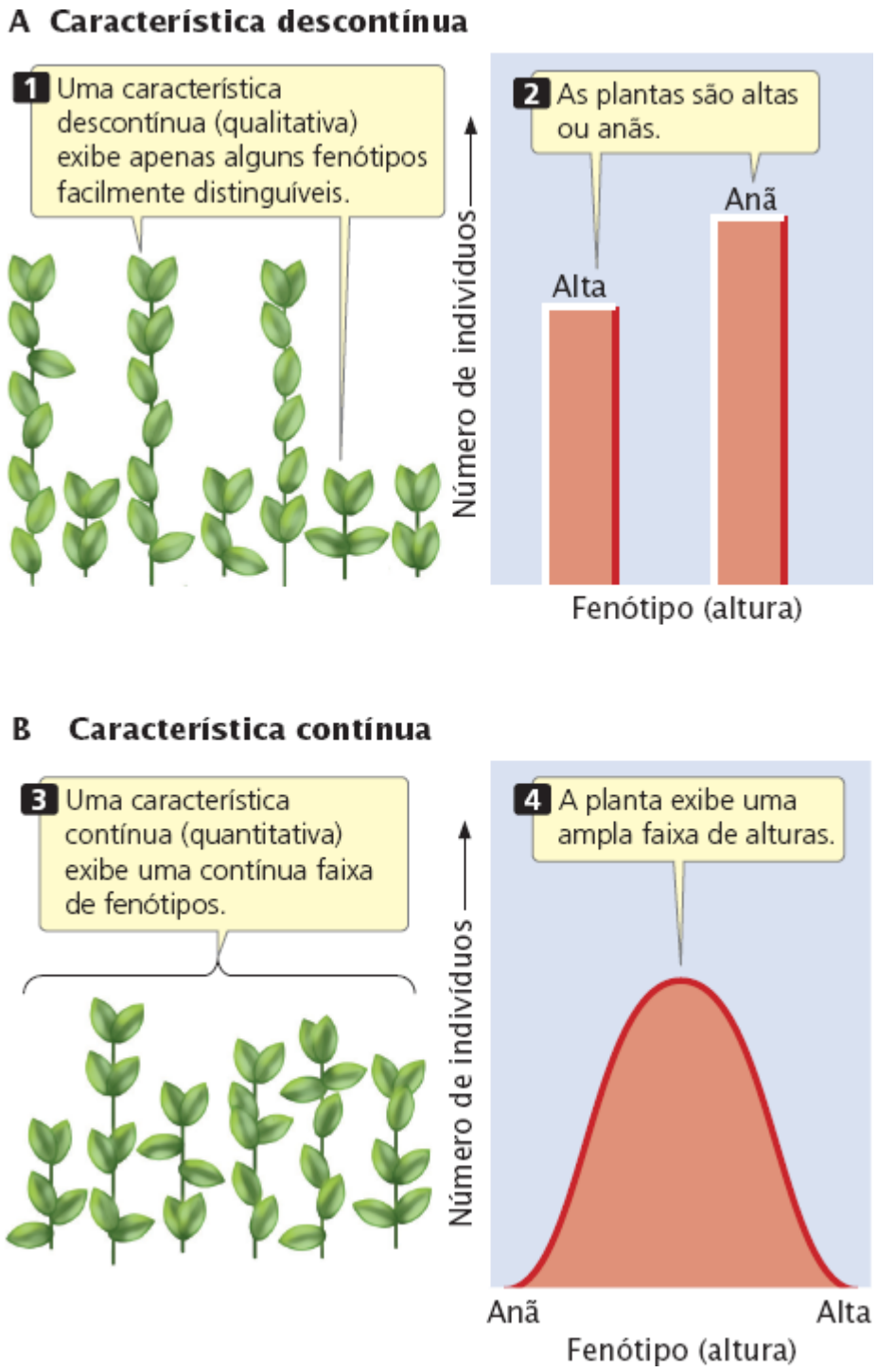


Figura 24.1 As características descontínuas e contínuas diferem no número de fenótipos exibidos.

As características quantitativas surgem a partir de dois fenômenos. Primeiro, muitas são poligênicas: são influenciadas por genes em vários *loci*. Se vários *loci* participarem, vários genótipos são possíveis, cada um produzindo um fenótipo

discretamente diferente. Segundo, as características quantitativas surgem quando os fatores ambientais afetam o fenótipo porque as diferenças ambientais resultam em um único genótipo produzindo uma gama de fenótipos. A maioria das características que variam continuamente é poligênica *e* influenciada por fatores ambientais e estas características são chamadas de multifatoriais.

Relação entre genótipo e fenótipo

Para algumas características descontínuas, a relação entre genótipo e fenótipo é direta. Cada genótipo produz um único fenótipo e a maioria dos fenótipos é codificada por um único genótipo. A dominância e a epistasia podem permitir que dois ou três genótipos produzam o mesmo fenótipo, mas a relação permanece simples. Esta relação simples entre genótipo e fenótipo permitiu que Mendel decifrasse as regras básicas da herança a partir de seus cruzamentos com as ervilhas; também permitiu que pudéssemos prever o resultado de cruzamentos genéticos e designar genótipos para os indivíduos.

Para as características quantitativas, a relação entre genótipo e fenótipo é, com frequência, mais complexa. Se a característica for poligênica, vários genótipos diferentes são possíveis, muitos dos quais podem produzir o mesmo fenótipo. Por exemplo, considere uma planta cuja altura é determinada por três *loci* (*A*, *B* e *C*), cada um com dois alelos. Pressuponha que um alelo em cada *locus* (A^+, B^+ e C^+) codifica um hormônio da planta que faz a planta crescer 1 cm acima da altura inicial de 10 cm. O outro alelo em cada *locus* (A^-, B^- e C^-) não codifica um hormônio da planta e então não contribui para a altura adicional. Se considerarmos apenas os dois alelos em um único *locus*, são possíveis 3 genótipos (A^+A^+, A^+A^- e A^-A^-). Se todos os três *loci* forem levados em consideração, existe um total de $3^3 = 27$ possíveis genótipos *multilocus* (A^+A^+ B^+B^+ C^+C^+, A^+A^- B^+B^+ C^+C^+ etc.). Embora existam 27 genótipos, eles produzem apenas sete genótipos (10, 11, 12, 13, 14, 15 e 16 cm de altura). Alguns dos fenótipos produzem o mesmo fenótipo (**Quadro 24.1**), por exemplo, os genótipos A^+A^- B^-B^- C^-C^-, A^-A^- B^+B^- C^-C^-, e A^-A^- B^-B^- C^+C^- têm um gene que codifica um hormônio da planta. Cada um destes genótipos produz uma dose de hormônio e resulta em uma planta de 11 cm de altura. Mesmo neste simples exemplo de apenas três *loci*, a relação entre genótipo e fenótipo é bem complexa. Quanto mais *loci* codificarem uma característica, maior será a complexidade. À medida que aumenta o número de *loci* que codificam uma característica, o número de potenciais fenótipos aumenta, e se torna mais difícil distinguir as diferenças entre fenótipos individuais.

A influência do ambiente em uma característica também pode complicar a relação entre genótipo e fenótipo. Por causa dos efeitos ambientais, o mesmo genótipo pode produzir uma variedade de potenciais fenótipos. As variações fenotípicas de diferentes genótipos podem se sobrepor, dificultando saber se os indivíduos têm diferentes fenótipos por causa das diferenças genéticas ou ambientais (**Figura 24.2**).

Em poucas palavras, a simples relação entre genótipo e fenótipo que existe para muitas características qualitativas (descontínuas) não existe nas características quantitativas, e é impossível indicar um genótipo de um indivíduo com base apenas no seu fenótipo. Os métodos usados para analisar as características qualitativas (examinar as razões fenotípicas de descendentes de um cruzamento genético) não funcionarão para as características quantitativas. Nosso objetivo permanece o mesmo: queremos fazer previsões sobre os fenótipos de descendentes de um cruzamento genético. Também queremos saber o quanto de variação de uma característica é o resultado de diferenças genéticas e o quanto é resultado de diferenças ambientais. Para responder a estas perguntas, devemos usar os métodos estatísticos que permitem que façamos previsões sobre a herança de fenótipos na ausência de informações sobre os genótipos envolvidos. **Resolva o Problema 16**

Quadro 24.1 Exemplo hipotético de altura de planta determinada pelos pares de alelos em cada um dos três *loci*.

Genótipo da planta	Doses de hormônio	Altura (cm)
A^-A^- B^-B^- C^-C^-	0	10
A^+A^- B^-B^- C^-C^- A^-A^- B^+B^- C^-C^- A^-A^- B^-B^- C^-C^+	1	11
A^+A^+ B^-B^- C^-C^- A^-A^- B^+B^+ C^-C^- A^-A^- B^-B^- C^+C^+ A^+A^- B^+B^- C^-C^- A^+A^- B^-B^- C^+C^- A^-A^- B^+B^- C^+C^-	2	12
A^+A^+ B^+B^- C^-C^- A^+A^- B^+B^+ C^-C^- A^+A^- B^-B^- C^+C^+ A^-A^- B^+B^+ C^+C^- A^-A^- B^+B^- C^+C^+ A^+A^- B^-B^- C^+C^-	3	13
A^+A^+ B^+B^+ C^-C^- A^+A^+ B^-B^- C^+C^- A^-A^- B^+B^+ C^+C^- A^+A^- B^+B^- C^+C^-	4	14
A^+A^+ B^+B^+ C^+C^- A^+A^+ B^+B^- C^+C^+ A^+A^- B^+B^+ C^+C^+	5	15
A^+A^+ B^+B^+ C^+C^+	6	16

$A\,A\;\;B\,B\;\;C\,C$ $A^+A^+\,B^-B^-\,C^+C^-$ $A^+A^-\,B^+B^+\,C^-C^-$ $A^-A^-\,B^+B^+\,C^+C^-$ $A^+A^-\,B^-B^-\,C^+C^+$ $A^-A^-\,B^+B^-\,C^+C^+$ $A^+A^-\,B^+B^-\,C^+C^-$	3	13
$A^+A^+\,B^+B^+\,C^-C^-$ $A^+A^+\,B^+B^-\,C^+C^-$ $A^+A^-\,B^+B^+\,C^+C^-$ $A^-A^-\,B^+B^+\,C^+C^+$ $A^+A^+\,B^-B^-\,C^+C^+$ $A^+A^-\,B^+B^-\,C^+C^+$	4	14
$A^+A^+\,B^+B^+\,C^+C^-$ $A^+A^-\,B^+B^+\,C^+C^+$ $A^+A^+\,B^+B^-\,C^+C^+$	5	15
$A^+A^+\,B^+B^+\,C^+C^+$	6	16

Nota: Cada alelo + contribui com 1 cm de altura acima do valor inicial de 10 cm.

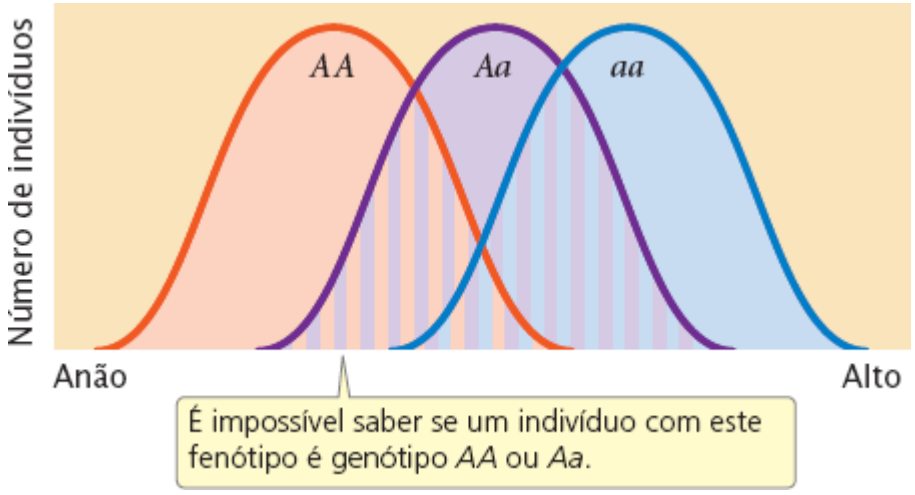


Figura 24.2 Para uma característica quantitativa, cada genótipo pode produzir uma gama de possíveis fenótipos. Neste exemplo hipotético, os fenótipos produzidos pelos genótipos *AA*, *Aa* e *aa* se sobrepõem.

Tipos de características quantitativas

Antes de examinarmos de perto as características poligênicas e os métodos estatísticos relevantes, precisamos definir com mais clareza o que significa uma característica quantitativa. Até aqui, analisamos apenas as características quantitativas que variam de forma contínua em uma população. Uma *característica contínua* pode teoricamente assumir qualquer valor entre dois extremos; o número de fenótipos é limitado apenas pela nossa capacidade de medir com precisão o fenótipo. A altura do humano é uma característica contínua porque, dentro de alguns limites, as pessoas podem teoricamente ter qualquer altura. Embora o número de fenótipos possíveis com uma característica contínua seja infinito, agrupamos fenótipos similares por conveniência; podemos dizer que duas pessoas têm 1,5 m, mas a medida cuidadosa pode mostrar que uma é ligeiramente mais alta que a outra.

Algumas características não são contínuas, mas são, contudo, consideradas contínuas porque são determinadas por múltiplos fatores genéticos e ambientais. As **características merísticas**, por exemplo, são medidas por números inteiros. Um exemplo é o tamanho da ninhada: um camundongo-fêmea pode ter 4, 5 ou 6 filhotes, mas não 4,13 filhotes. Uma característica merística tem um número limitado de fenótipos distintos, mas a determinação envolvida da característica ainda pode ser quantitativa. Estas características devem, portanto, ser analisadas com as mesmas técnicas que usamos para estudar as características quantitativas

contínuas.

Outro tipo de característica quantitativa é uma **característica com limiar**, que está simplesmente presente ou ausente. Embora as características com limiar exibam apenas dois fenótipos, elas são consideradas quantitativas porque também são determinadas por múltiplos fatores genéticos e ambientais. A expressão da característica depende de uma suscetibilidade envolvida (em geral chamada de tendência ou risco) que varia continuamente. Quando a suscetibilidade é maior que um valor limiar, um traço específico é expresso (**Figura 24.3**). As doenças são características com limiar porque muitos fatores, genéticos e ambientais, contribuem para sua suscetibilidade. Se estiverem presentes fatores de suscetibilidade suficientes, a doença se desenvolve; do contrário, ela está ausente. Embora nos voltemos para a genética das características contínuas neste capítulo, os mesmos princípios se aplicam a muitas características merísticas e com limiar.

O fato de uma característica poder ser medida em uma escala contínua não significa que ela exiba variação quantitativa. Uma das características estudadas por Mendel era a altura da planta da ervilha, que pode ser descrita ao medir o comprimento do caule da planta. Entretanto, as plantas específicas de Mendel exibiam apenas dois fenótipos distintos (algumas eram altas e outras eram baixas) e estas diferenças eram determinadas por alelos em um único *locus*. As diferenças que Mendel estudou eram naturalmente descontínuas.

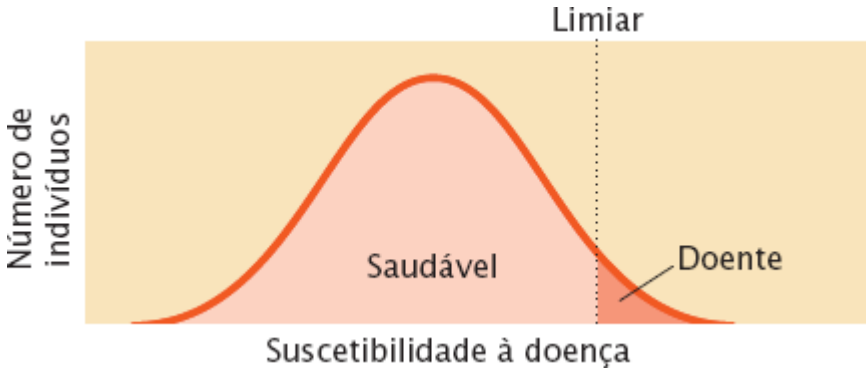


Figura 24.3 As características com limiar apresentam dois fenótipos possíveis, o traço está presente ou ausente, mas elas são quantitativas porque a suscetibilidade envolvida à característica varia continuamente. Quando a suscetibilidade excede um valor limiar, a característica é expressa.

Conceitos

As características para as quais os fenótipos podem variar continuamente são chamadas de características quantitativas. Para a maioria das características quantitativas, a relação entre genótipo e fenótipo é complexa. Algumas características para as quais os fenótipos não variam continuamente também são consideradas quantitativas porque são influenciadas por múltiplos genes e fatores ambientais.

Herança poligênica

Após a redescoberta do trabalho de Mendel em 1900, surgiram perguntas sobre a herança de características que variam continuamente. Estas características já foram o foco de um grupo de biólogos e estatísticos, liderados por Francis Galton, que usou procedimentos estatísticos para examinar a herança de características quantitativas como a altura e a inteligência humanas. Os resultados destes estudos mostraram que as características quantitativas eram pelo menos parcialmente herdadas, embora não se conhecesse ainda o mecanismo de herança. Alguns especialistas em biometria argumentaram que as características quantitativas não poderiam ser explicadas pelos princípios mendelianos, enquanto outros acreditavam que os princípios de Mendel que atuam em numerosos genes (poligenes) poderiam justificar a herança das características quantitativas.

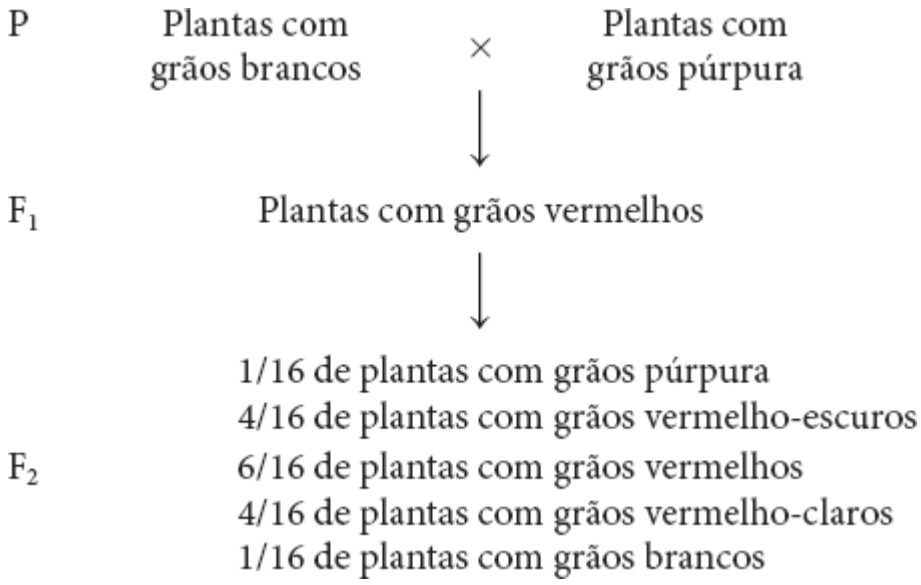
Este conflito começou a ser resolvido com o trabalho de Wilhelm Johannsen, que mostrou que a variação contínua no peso dos feijões era influenciada por fatores genéticos e ambientais. George Udny Yule, um matemático, propôs, em 1906, que vários genes que atuam juntos poderiam produzir características contínuas. Esta hipótese foi posteriormente confirmada por Herman Nilsson-Ehle, que trabalhava com trigo e tabaco, e por Edward East, que trabalhava com milho. O argumento foi finalmente abandonado em 1918, quando Ronald Fisher demonstrou que a herança de características quantitativas poderia ser explicada pelos efeitos cumulativos de muitos genes, cada um seguindo as regras de Mendel.

Cor do grão do trigo

Para ilustrar como múltiplos genes que atuam em uma característica podem produzir uma faixa contínua de fenótipos, vamos examinar uma das primeiras demonstrações da herança poligênica. Nilsson-Ehle estudou a cor do grão do trigo e descobriu que

a intensidade da pigmentação vermelha era determinada por três *loci* não ligados, cada um com dois alelos.

Cruzamento de Nilsson-Ehle. Nilsson-Ehle obteve variedades homozigotas de trigo com cores diferentes. Como Mendel, ele fez cruzamentos entre estas variedades homozigotas e estudou as razões de fenótipos nos descendentes. Em um experimento, ele cruzou uma variedade de trigo com grãos brancos com uma variedade de grãos púrpura (vermelho muito escuro) e obteve os seguintes resultados:



Interpretação do cruzamento. Nilsson-Ehle interpretou esta razão fenotípica como resultado da segregação dos alelos em dois *loci* (embora ele tenha encontrado alelos em três *loci* que afetam a cor do grão, as duas variedades usadas neste cruzamento tinham diferenças em apenas dois dos *loci*). Ele propôs que existiam dois alelos em cada *locus*: um que produzia o pigmento vermelho e outro que não produzia pigmento. Indicamos os alelos que codificam pigmento como A^+ *e* B^+ e os alelos que não codificam pigmento como A^- e B^-. Nilsson-Ehle reconheceu que os efeitos dos genes eram aditivos. Cada gene parecia contribuir igualmente para a cor; então o genótipo geral poderia ser determinado ao adicionar os efeitos de todos os genes, como descrito na tabela a seguir.

Genótipo	Doses de pigmento	Fenótipo
$A^+A^+\,B^+B^+$	4	Púrpura
$A^+A^+\,B^+B^-$ $A^+A^-\,B^+B^+$	3	Vermelho-escuro
$A^+A^+\,B^-B^-$ $A^-A^-\,B^+B^+$ $A^+A^-\,B^+B^-$	2	Vermelho
$A^+A^-\,B^-B^-$ $A^-A^-\,B^+B^-$	1	Vermelho-claro
$A^-A^-\,B^-B^-$	0	Branco

Observe que os fenótipos púrpura e branco são cada um codificado por um único genótipo, mas os outros fenótipos podem surgir a partir de vários genótipos diferentes.

A partir destes resultados, observamos que cinco fenótipos são possíveis quando os alelos em dois *loci* influenciam o fenótipo e os efeitos dos genes são aditivos. Quando os alelos em mais de dois *loci* influenciam o fenótipo, mais fenótipos são possíveis e a cor variaria continuamente entre branco e púrpura. Se os fatores ambientais influenciam a característica, os indivíduos do mesmo genótipo variariam na cor, dificultando mais ainda distinguir entre discretas classes fenotípicas. Felizmente, o ambiente tem pouca participação na determinação da cor do grão nos cruzamentos de Nilsson-Ehle e apenas alguns *loci* codificam a cor; então ele foi capaz de distinguir entre as diferentes classes fenotípicas. Esta capacidade permitiu que ele observasse a natureza mendeliana da característica.

Agora vamos ver como os princípios de Mendel explicam a razão obtida por Nilsson-Ehle nos seus descendentes F_2. Lembre-

se de que ele cruzou a variedade púrpura homozigota ($A^+A^+\ B^+B^+$) com a variedade branca homozigota ($A^-A^-\ B^-B^-$), produzindo os descendentes F_1 heterozigotos nos dois *loci* ($A^+A^-\ B^+B^-$). Este cruzamento é um cruzamento di-híbrido, como os que discutimos no Capítulo 3, exceto que ambos *loci* codificam o mesmo traço. Todas as plantas F_1 tinham dois alelos que produziam pigmento que permitiam duas doses de cor para criar os grãos vermelhos. Os tipos e proporções de descendentes esperados em F_1 podem ser encontrados ao aplicar os princípios de segregação e de segregação independente de Mendel.

Primeiro, vamos examinar os efeitos de cada *locus* separadamente. No primeiro *locus*, dois F_1 heterozigotos são cruzados ($A^+A^- \times A^+A^-$). Como aprendemos no Capítulo 3, quando dois heterozigotos são cruzados, esperamos os descendentes nas proporções 1/4 A^+A^+, 1/2 A^+A^- e 1/4 A^-A^-. No segundo *locus*, dois heterozigotos também são cruzados e, mais uma vez, esperamos os descendentes nas proporções de 1/4 B^+B^+, 1/2 B^+B^- e 1/4 B^-B^-.

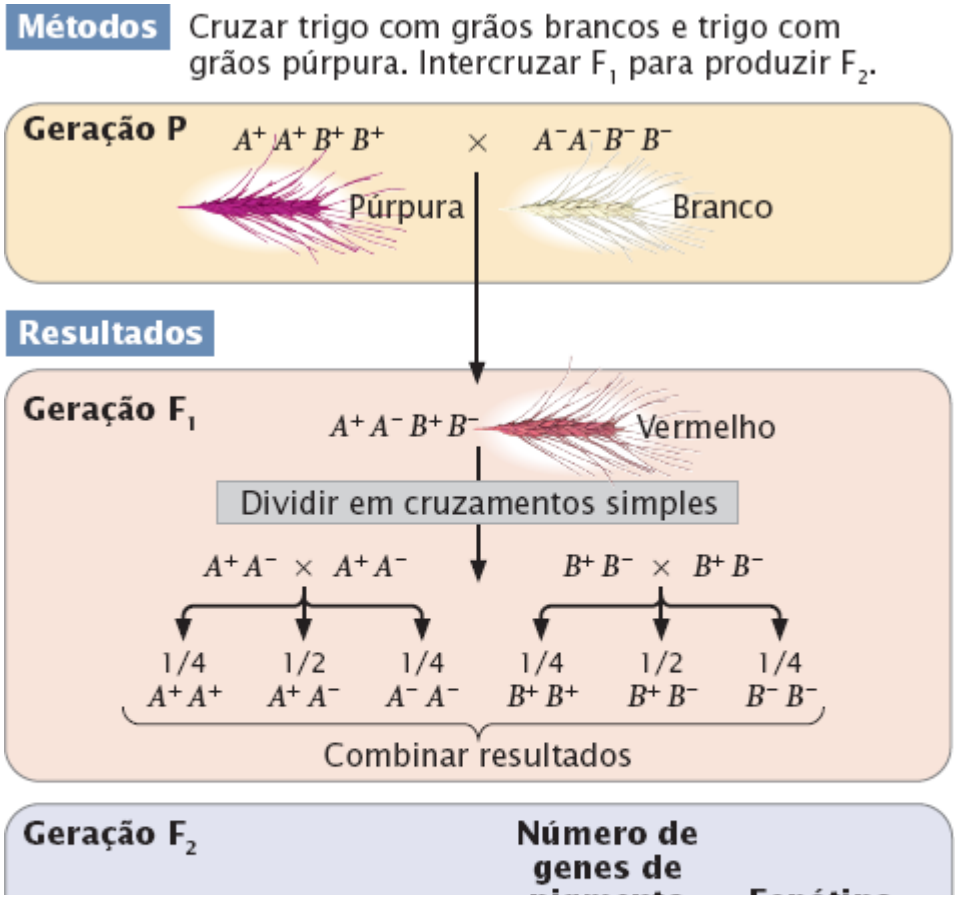
Para obter a probabilidade de combinações dos genes em ambos *loci*, é preciso usar a regra de multiplicação de probabilidade (ver *Capítulo 3*), que assume o princípio de segregação independente de Mendel. A proporção esperada de descendentes F_2 com o genótipo $A^+A^+\ B^+B^+$ é o produto da probabilidade de obter o genótipo A^+A^+ (1/4) e a probabilidade de obter o genótipo B^+B^+ (1/4) ou 1/4 × 1/4 = 1/16 (**Figura 24.4**). As probabilidades de cada um dos fenótipos podem ser obtidas ao adicionar as probabilidades de todos os genótipos que produzem aquele fenótipo. Por exemplo, o fenótipo vermelho é produzido por três genótipos:

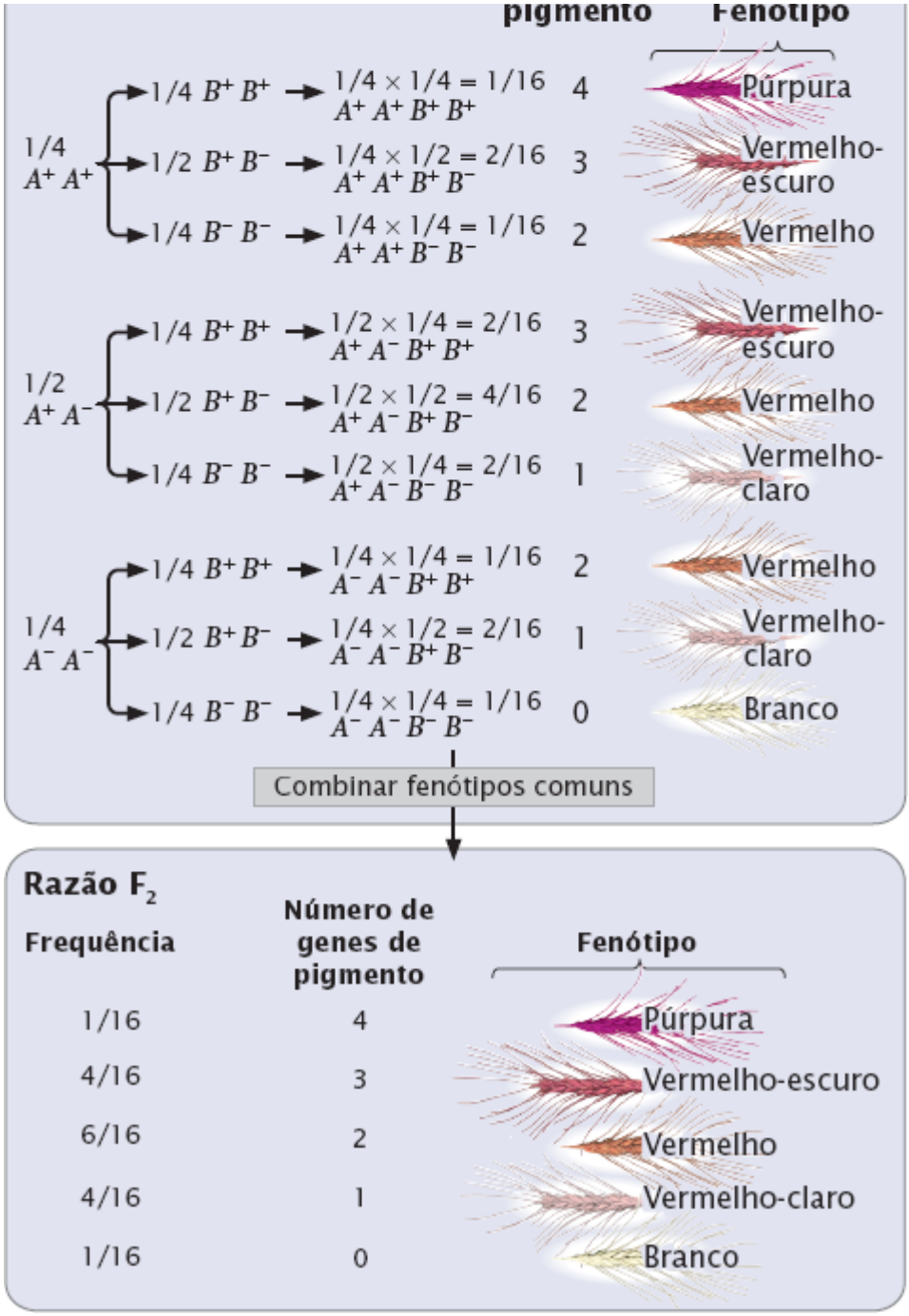
Genótipo	***Fenótipo***	***Probabilidade***
$A^+A^+\ B^-B^-$	Vermelho	1/16
$A^-A^-\ B^+B^+$	Vermelho	1/16
$A^+A^-\ B^+B^-$	Vermelho	1/4

Assim, a probabilidade geral de obter grãos vermelhos nos descendentes F_2 é 1/16 + 1/16 + 1/4 = 6/16. A Figura 24.4 mostra que a razão fenotípica esperada em F_2 é 1/16 púrpura, 4/16 vermelho-escuro, 4/16 vermelho, 4/16 vermelho-claro e 1/16 branco. Esta razão fenotípica é exatamente o que Nilsson-Ehle observou nos seus descendentes F_2, demonstrando que a herança de uma característica que varia continuamente como a cor do grão está de fato de acordo com os princípios básicos de Mendel.

Experimento

Pergunta: Como é herdado um traço contínuo, como a cor do grão do trigo?





Conclusão: A cor do grão no trigo é herdada de acordo com os princípios de Mendel ao atuar nos alelos em dois *loci*.

Figura 24.4 Nilsson-Ehle demonstrou que a cor do grão no trigo é herdada de acordo com os princípios de Mendel. A razão de fenótipos em F_2 pode ser determinada ao romper o cruzamento di-híbrido em dois cruzamentos simples de único *locus* e combinar os resultados usando a regra de multiplicação.

Conclusões e implicações. Os cruzamentos de Nilsson-Ehle demonstraram que a diferença entre a herança de genes que influenciam as características quantitativas e a herança de genes que influenciam as características descontínuas estão no número de *loci* que determinam a característica. Quando múltiplos *loci* afetam uma característica, mais genótipos são possíveis, então a relação entre o genótipo e fenótipo é menos evidente. À medida que o número de *loci* que afeta uma característica aumenta, o número de classes fenotípicas em F_2 aumenta (**Figura 24.5**).

Várias condições dos cruzamentos de Nilsson-Ehle simplificaram muito a herança poligênica da cor do grão e possibilitaram que ele identificasse a natureza mendeliana da característica. Primeiro, os genes que afetam a cor se separam em apenas dois ou três *loci*. Se os genes em vários *loci* estão separados, ele teria dificuldade em diferenciar as classes fenotípicas. Segundo, os genes que afetam a cor do grão têm efeitos rigorosamente aditivos, simplificando a relação entre genótipo e fenótipo. Terceiro, o ambiente quase não influencia o fenótipo; se os fatores ambientais modificassem os fenótipos, seria difícil distinguir entre cinco classes fenotípicas. Finalmente, os *loci* que Nilsson-Ehle estudou não estavam ligados, então os genes se segregavam

independentemente. Nilsson-Ehle foi sortudo: para muitas características poligênicas, estas condições simplificadas não estão presentes e a herança mendeliana destas características não é óbvia. **Resolva o Problema 17**

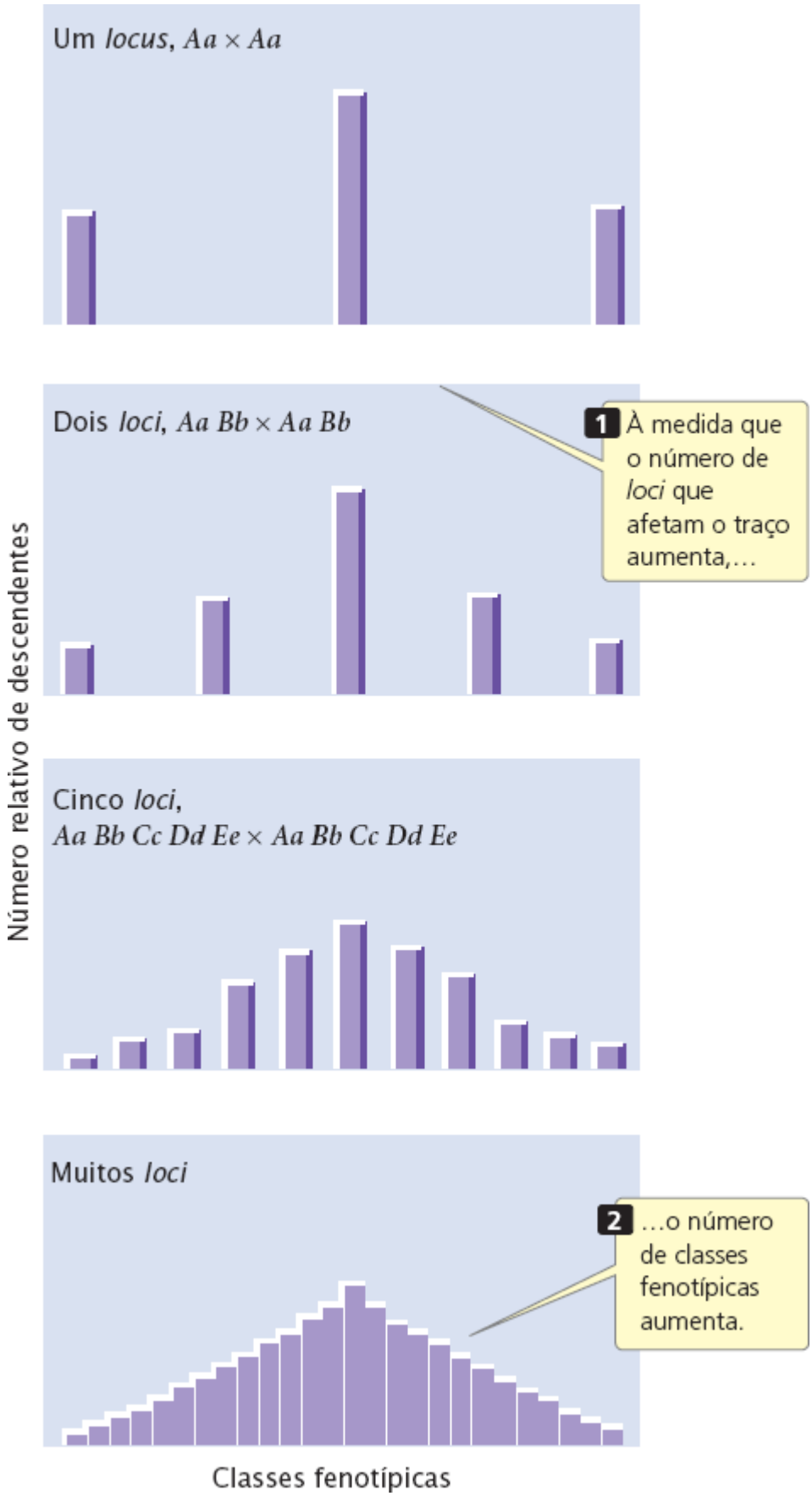


Figura 24.5 Os resultados do cruzamento indivíduos heterozigotos para diferentes números de *loci* que afetam uma característica.

Como determinar o número de genes para uma característica poligênica

A proporção de indivíduos F_2 que lembram um dos genitores originais pode ser usada para estimar o número de genes que

afetam um traço poligênico. Quando dois indivíduos homozigotos para diferentes alelos em um único *locus* são cruzados ($A^1\,A^1 \times A^2\,A^2$) e os F_1 resultantes são intercruzados ($A^1\,A^2 \times A^1\,A^2$), um quarto de F_2 deve ser homozigoto como cada um dos genitores originais. Se os genitores originais forem homozigotos para diferentes alelos em dois *loci*, como eram nos cruzamentos de Nilsson-Ehle, então $1/4 \times 1/4 = 1/16$ de F_2 lembrará um dos genitores homozigotos originais. Em geral, $(1/4)^n$ será a proporção de indivíduos nos descendentes F_2 que devem lembrar cada um dos genitores homozigotos originais, enquanto *n* é igual ao número de *loci* com um par segregante de alelos que afetam a característica. Esta equação nos fornece meios possíveis para determinar o número de *loci* que influenciam uma característica quantitativa.

Para ilustrar o uso desta equação, considere que cruzamos duas variedades homozigotas distintas de ervilhas com diferença de 16 cm na altura, intercruzamos F_1 e descobrimos que aproximadamente 1/256 de F_2 são semelhantes a uma das variedades genitoras homozigotas originais. Este desfecho sugeriria que quatro *loci* com pares de alelos segregantes ($1/256 = (1/4)^4$) são responsáveis pela diferença de altura entre as duas variedades. Como as duas cepas homozigotas apresentam uma diferença de 16 cm na altura e existem quatro *loci* e cada qual tem dois alelos (oito alelos no total), cada um dos alelos contribui com 16 cm/8 = 2 cm na altura.

Este método para determinar o número de *loci* que afetam diferenças fenotípicas requer o uso de cepas homozigotas, o que pode ser difícil de se obter em alguns organismos. Ele também considera que todos os genes que influenciam a característica têm efeitos iguais, que seus efeitos são aditivos e que os *loci* não estão ligados. Para as características poligênicas, estas considerações não são válidas, e então este método de determinar o número de genes que afetam uma característica tem aplicação limitada. **Resolva o Problema 19**

Conceitos

Os mesmos princípios determinam a herança de características quantitativas e descontínuas, contudo, mais genes participam na determinação das características quantitativas.

Checagem dos conceitos 1

Explique resumidamente como o número de genes que influencia um traço poligênico pode ser determinado.

24.2 Métodos estatísticos são cruciais para analisar características quantitativas

Uma vez que as características quantitativas são descritas como uma medida e são influenciadas por múltiplos fatores, sua herança tem de ser analisada sob a ótica da estatística. Esta seção vai explicar os conceitos básicos de estatística usados para analisar as características quantitativas.

Distribuições

Compreender a base genética de qualquer característica começa com uma descrição dos números e tipos de fenótipos presentes em um grupo de indivíduos. A variação fenotípica em um grupo pode ser representada de forma conveniente por uma **distribuição de frequência**, que é um gráfico das frequências (número de proporções) de diferentes fenótipos (**Figura 24.6**). Em uma típica distribuição de frequências, as classes de fenótipos são inseridas no eixo horizontal (*x*) e os números (ou proporções) de indivíduos em cada classe são inseridos no eixo vertical (*y*). Uma distribuição de frequência é um método conciso para resumir todos os fenótipos de uma característica quantitativa.

Conectar os pontos de uma distribuição de frequência com uma linha cria uma curva característica da distribuição. Várias características quantitativas exibem uma curva simétrica (em forma de sino) chamada de **distribuição normal** (**Figura 24.7 A**). As distribuições normais surgem quando vários fatores independentes contribuem para uma medida, como é o caso nas características quantitativas. Dois outros tipos comuns de distribuições (distorcida e bimodal) estão ilustrados na **Figura 24.7 B** e **C**.

Amostras e populações

Frequentemente, os biólogos precisam descrever a distribuição dos fenótipos exibidos por um grupo de indivíduos. Podemos descrever a altura dos estudantes da University of Texas (UT), mas a UT tem mais de 40.000 alunos e não seria prático medir a todos. Os cientistas estão constantemente lidando com este problema: o grupo de interesse, chamado de **população**, é muito grande para um censo completo. Uma solução é medir uma coleção menor de indivíduos, chamada de **amostra**, e usar estas medidas para descrever a população.

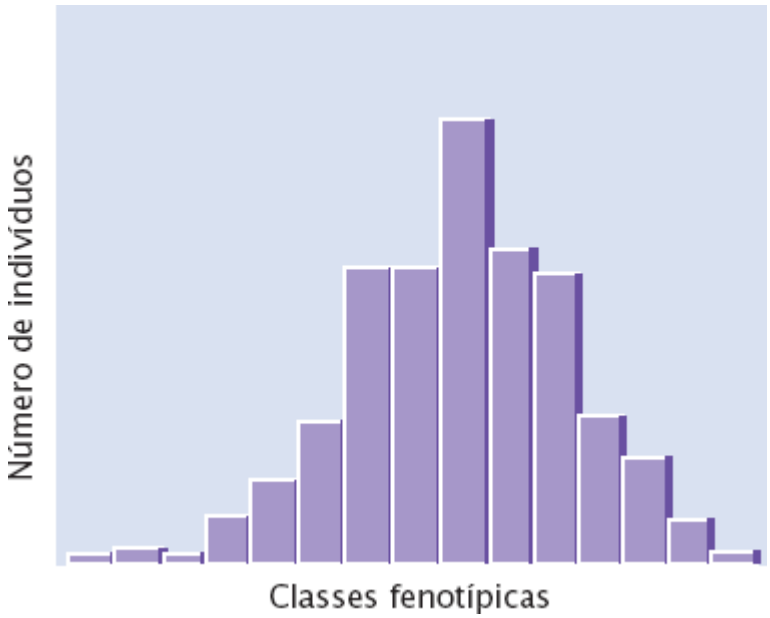
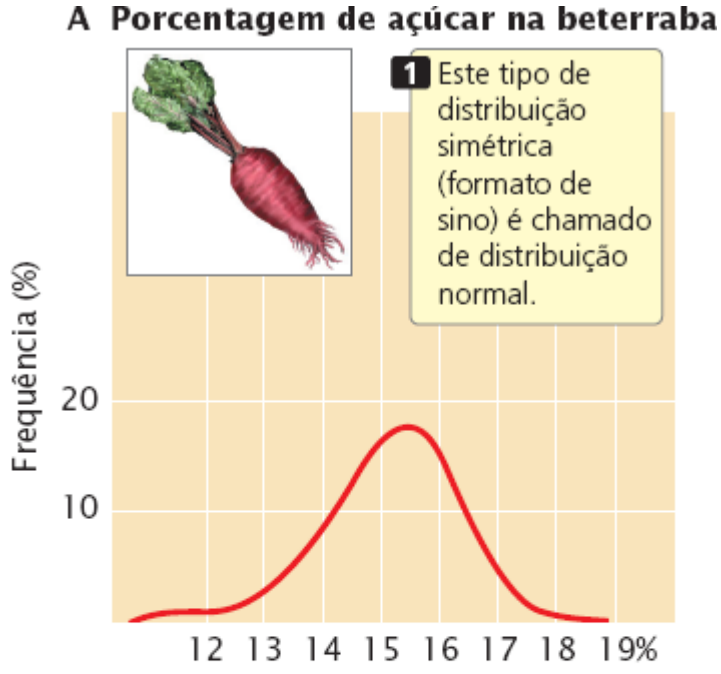


Figura 24.6 Uma distribuição de frequência é um gráfico que apresenta o número ou a proporção de diferentes fenótipos. Os valores fenotípicos são inseridos no eixo horizontal e os números (ou proporções) de indivíduos de cada classe são inseridos no eixo vertical.

Para fornecer uma descrição exata da população, uma boa amostra precisa ter várias características. Primeiro, tem de ser representativa da população inteira, por exemplo, se nossa amostra consistir inteiramente em membros da equipe de basquete da UT, estaríamos superestimando a altura verdadeira dos estudantes. Uma forma para garantir que uma amostra seja representativa da população é selecionar os membros da amostra aleatoriamente. Segundo, a amostra deve ser grande o suficiente de modo que as diferenças de probabilidade entre os indivíduos na amostra e na população geral não distorçam a estimativa das medidas da população. Se medirmos apenas três estudantes da UT e, por acaso, todos forem baixos, estaríamos subestimando a verdadeira altura da população de estudantes. A estatística pode fornecer informações sobre o quanto confiar nas estimativas com base em amostras aleatórias.



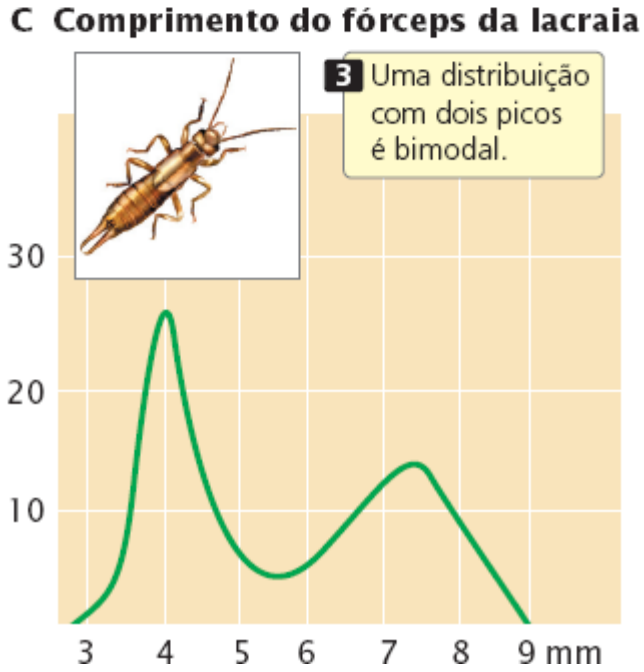


Figura 24.7 **As distribuições de fenótipos podem adotar vários formatos diferentes.**

Conceitos

Na estatística, a população é o grupo de interesse, uma amostra é um subconjunto da população. A amostra deve ser representativa da população e grande o suficiente para minimizar as diferenças de probabilidade entre a população e a amostra.

Checagem dos conceitos 2

Uma geneticista quer saber se a asma é causada por uma mutação no gene *DS112*. Ela coleta DNA de 120 pessoas com asma e de 100 pessoas saudáveis e sequencia o DNA. Ela descobre que 35 das pessoas com asma e nenhuma das saudáveis têm uma mutação no gene *DS112*. Qual é a população neste estudo?

a. As 120 pessoas com asma.
b. As 100 pessoas saudáveis.
c. As 35 pessoas com uma mutação no seu gene.
d. Todas as pessoas com asma.

Média

A **média**, também chamada de média aritmética, fornece informações sobre o centro da distribuição. Se medirmos a altura de garotos de 10 e 18 anos e organizarmos uma distribuição de frequência para cada grupo, descobriríamos que ambas as

distribuições são normais, mas as duas distribuições estariam centralizadas em diferentes alturas e esta diferença estaria indicada nas suas diferentes médias (**Figura 24.8**).

Suponha que temos cinco medidas de altura em centímetros: 160, 161, 167, 164 e 165. Se representamos um grupo de medidas como x_1, x_2, x_3… então a média ($\bar{x}$) é calculada ao adicionar todas as medidas individuais e dividir pelo número total de medidas na amostra (n):

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ \ldots \ + x_n}{n} \tag{24.1}$$

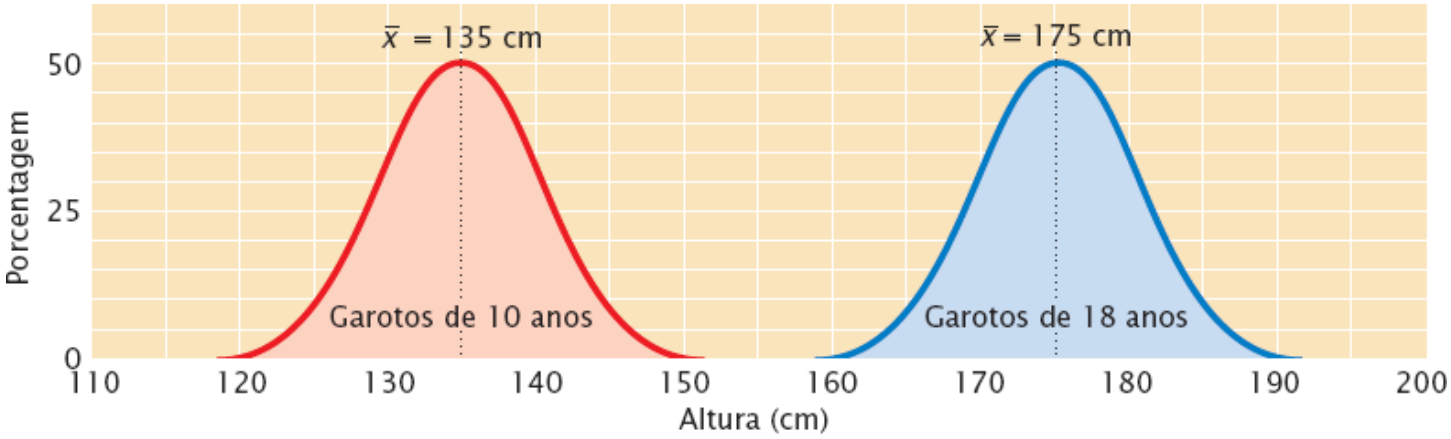


Figura 24.8 A média informa o centro de uma distribuição. Ambas as distribuições de alturas de garotos de 10 e 18 anos são normais, mas têm diferentes localizações ao longo de uma sequência contínua de altura, o que torna suas médias diferentes.

No nosso exemplo, x_1 = 160, x_2 = 161, x_3 = 167, e daí em diante. A altura média ($\bar{x}$) é igual:

$$\bar{x} = \frac{160 + 161 + 167 + 164 + 165}{5} = \frac{817}{5} = 163{,}4$$

Uma forma simplificada para representar esta fórmula é:

$$\bar{x} = \frac{\sum x_i}{n} \tag{24.2}$$

ou

$$\bar{x} = \frac{1}{n}\sum x_i \tag{24.3}$$

em que o símbolo Σ significa “a soma de” e x_i representa os valores individuais de x.

Variância e desvio padrão

Uma medida estatística que fornece informações importantes sobre uma distribuição é a **variância**, que indica a variabilidade de um grupo de medidas, ou como a distribuição está espalhada. As distribuições podem ter a mesma média, mas diferentes variâncias (**Figura 24.9**). Quanto maior a variância, maior o espalhamento das medidas em uma distribuição ao redor de sua média.

A variância (s^2) é definida como o desvio médio ao quadrado a partir da média:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} \tag{24.4}$$

Para calcular a variância, (1) subtraímos a média de cada medida e o quadrado do valor obtido, (2) adicionamos todos os desvios elevados ao quadrado e (3) dividimos esta soma pelo número de medidas originais menos 1.

Outra medida estatística que está intimamente relacionada com a variância é o **desvio padrão** (s), que é definido como a raiz quadrada da variância:

$$s = \sqrt{s^2} \tag{24.5}$$

Enquanto a variância é expressa em unidades elevadas ao quadrado, o desvio padrão está na mesma unidade que as medidas originais; então o desvio padrão é preferido para descrever a variabilidade de uma medida.

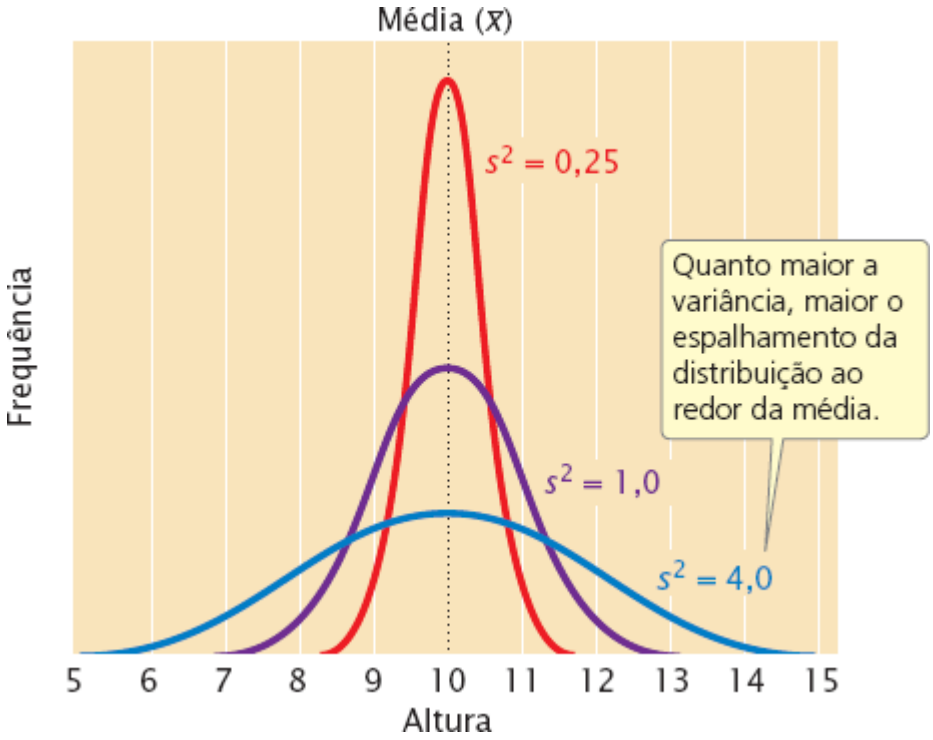


Figura 24.9 A variância informa sobre a variabilidade de um grupo de fenótipos. Aqui estão três distribuições com a mesma média, mas diferentes variâncias.

Uma distribuição normal é simétrica, então a média e o desvio padrão são suficientes para descrever seu formato. A média mais ou menos um desvio padrão ($\bar{x} \pm s$) inclui aproximadamente 66% das medidas em uma distribuição normal, a média mais ou menos dois desvios padrões ($\bar{x} \pm 2$ s) inclui aproximadamente 95% das medidas e a média mais ou menos três desvios padrões ($\bar{x} \pm 3$ s) inclui aproximadamente 99% das medidas (**Figura 24.10**). Assim, apenas 1% de uma população normalmente distribuída está fora da faixa de ($\bar{x} \pm 3$ s). Resolva o Problema 22

Conceitos

A média e a variância descrevem uma distribuição de medidas: a média informa a localização do centro de uma distribuição e a variância informa sua variabilidade.

Checagem dos conceitos 3

As medidas de uma distribuição com _____ maior serão mais espalhadas:

a. Média.
b. Variância.
c. Desvio padrão.
d. Variância e desvio padrão.

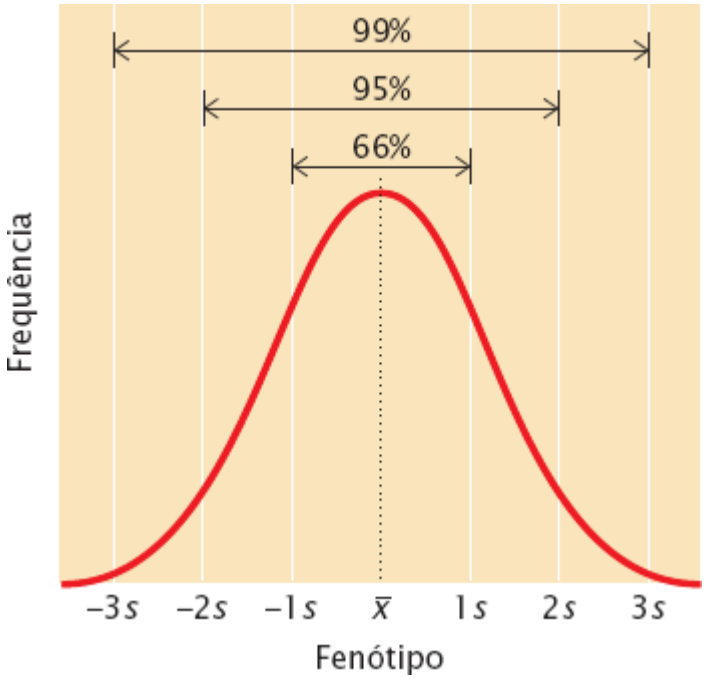


Figura 24.10 As proporções de uma distribuição normal ocupadas por mais ou menos um, dois ou três desvios padrões da média.

Correlação

A média e a variância podem ser usadas para descrever uma característica individual, mas os geneticistas estão interessados em mais de uma característica. Frequentemente, duas ou mais características variam juntas. Por exemplo, o número e o peso dos ovos produzidos por galinhas são importantes na indústria avícola. Estas duas características não são independentes uma da outra. Existe uma relação inversa entre o número de ovos e seu peso: as galinhas que põem mais ovos produzem ovos menores. Este tipo de relação entre duas características é chamada **correlação**. Quando duas características estão correlacionadas, é provável que mudança em uma delas esteja associada com mudança na outra.

As correlações entre características são medidas por um **coeficiente de correlação** (indicado como *r*), que mede a força de sua associação. Considere duas características, como a altura do humano (*x*) e comprimento do braço (*y*). Para determinar como estas características estão correlacionadas, primeiro obtemos a covariância (*cov*) de *x* e *y*:

$$\mathrm{cov}_{xy} = \frac{\sum (x_i - \bar{x})\,(y_i - \bar{y})}{n - 1} \tag{24.6}$$

A covariância é calculada ao (1) tomar um valor *x* para um indivíduo e subtraí-lo da média de $x\,\bar{x}$; (2) tomar o valor *y* do mesmo indivíduo e subtraí-lo da média de $y\,\bar{y}$; (3) multiplicar os resultados destas duas subtrações: (4) adicionar os resultados para todos os pares *xy*; e (5) dividir esta soma por $n - 1$ (em que *n* é igual ao número de pares *xy*).

O coeficiente de correlação (*r*) é obtido ao dividir a covariância de *x* e *y* pelo produto dos desvios padrões de *x* e *y*:

$$r = \frac{\mathrm{cov}_{xy}}{s_x s_y} \tag{24.7}$$

Um coeficiente de correlação pode variar teoricamente de –1 a +1. Um valor positivo indica que existe uma associação direta entre as variáveis (**Figura 24.11 A**): à medida que uma variável aumenta, a outra tende a aumentar. Existe uma correlação positiva para altura e peso nos humanos: as pessoas altas tendem a pesar mais. Um coeficiente de correlação negativo indica que existe uma relação inversa entre as duas variáveis (**Figura 24.11 B**): à medida que uma variável aumenta, a outra tende a diminuir (como é o caso do número de ovos e o peso do ovo nas galinhas).

O valor absoluto do coeficiente de correlação (o tamanho do coeficiente, ignorando seu sinal) fornece informações sobre a extensão da associação entre as variáveis. Um coeficiente de –1 ou +1 indica uma perfeita correlação entre as variáveis, o que significa que uma mudança em *x* está sempre acompanhada por uma mudança proporcional em *y*. Os coeficientes de correlação próximos de –1 ou +1 indicam uma forte associação entre as variáveis: uma mudança em *x* quase sempre está associada com um aumento proporcional em *y*, como observado na **Figura 24.11 C**. Por outro lado, um coeficiente de correlação próximo de 0

indica uma correlação fraca: uma mudança em *x* está associada com uma mudança em *y*, mas não sempre (**Figura 24.11 D**). Uma correlação de 0 indica que não existe associação entre as variáveis (**Figura 24.11 E**).

Um coeficiente de correlação pode ser calculado para duas variáveis medidas para o mesmo individuo, como a altura (*x*) e o peso (*y*). Um coeficiente de correlação também pode ser calculado para uma única variável medida para pares de indivíduos. Por exemplo, podemos calcular para o peixe a correlação entre o número de vértebras de um genitor (*x*) e o número de vértebras de seus descendentes (*y*), como apresentado na **Figura 24.12**. Esta abordagem é usada com frequência na genética quantitativa.

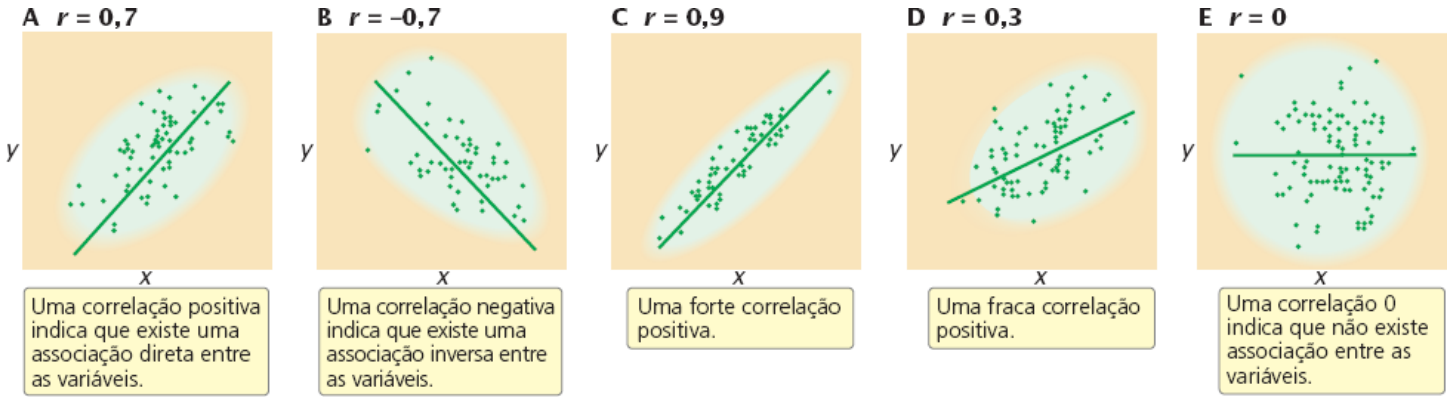


Figura 24.11 O coeficiente de correlação descreve a relação entre duas ou mais variáveis.

Uma correlação entre as duas variáveis indica apenas que elas estão associadas; ela não implica uma relação de causa e efeito. A correlação também não significa que os valores de duas variáveis são os mesmos; ele significa apenas que uma mudança em uma variável está associada com uma mudança proporcional na outra variável. Por exemplo, as variáveis *x* e *y* na lista a seguir estão quase perfeitamente correlacionadas, com um coeficiente de correlação de 0,99.

	Valor de *x*	Valor de *y*
	12	123
	14	140
	10	110
	6	61
	3	32
Média aritmética:	9	90

Figura 24.12 Um coeficiente de correla ção pode ser calculado para uma única variável medida para pares de indivíduos. Aqui, os números de vértebras da genitora e dos descendentes do peixe *Zoarces viviparus* são comparados.

É encontrada uma alta correlação entre estas variáveis *x* e *y*; valores maiores de *x* estão sempre associados com valores maiores de *y*. Observe que os valores de *y* são 10 vezes maiores que os valores de *x* correspondentes, então, embora *x* e *y* estejam correlacionados, não são idênticos. A distinção entre correlação e identidade se torna importante quando consideramos os efeitos da hereditariedade e ambiente sobre a correlação das características. Resolva o Problema 25

Regressão

A correlação informa apenas a força e a direção da associação entre as variáveis. Entretanto, queremos saber mais do que duas variáveis que estão associadas; queremos prever o valor de uma variável considerando o valor de outra.

Existe uma correlação positiva entre o peso do corpo dos pais e o peso do corpo dos seus descendentes; esta correlação existe em parte porque os genes influenciam o peso do corpo e os pais e as crianças têm os genes em comum. Graças a esta associação entre fenótipos de pais e descendentes, podemos prever o peso de um indivíduo com base no peso dos seus pais. Este tipo de previsão estatística é chamada de **regressão**. Esta técnica é importante na genética quantitativa porque permite que possamos prever as características dos descendentes de um dado cruzamento, mesmo sem saber os genótipos que codificam as características.

A regressão pode ser compreendida ao inserir uma série de valores de *x* e *y*. A **Figura 24.13** ilustra a relação entre o peso do pai (*x*) e o peso de seu filho (*y*). Cada par pai-filho é representado por um ponto no gráfico. A relação geral entre estas duas variáveis é descrita pela linha de regressão, que é a linha que melhor encaixa todos os pontos no gráfico (os desvios dos pontos a partir da linha são minimizados). A linha de regressão define a relação entre as variáveis *x* e y e pode ser representada por

$$y = a + bx \tag{24.8}$$

Na Equação 24.8, *x* e *y* representam as variáveis *x* e *y* (neste caso, o peso do pai e do filho, respectivamente). A variável *a* é o intercepto *y* da linha, que é o valor esperado de *y* quando *x* é 0. A variável *b* é a inclinação da linha de regressão, também chamada de **coeficiente de regressão**.

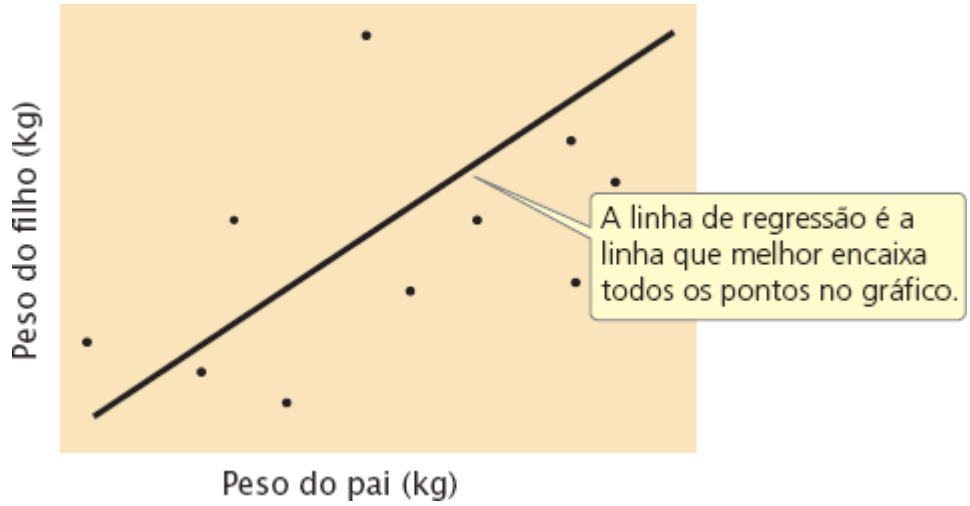


Figura 24.13 Uma linha de regressão define a relação entre duas variáveis. Aqui está ilustrada uma regressão dos pesos dos pais contra os pesos dos filhos. Cada par pai-filho é representado por um ponto no gráfico: o valor *x* de um ponto é o peso do pai e o valor *y* do ponto é o peso do filho.

Tentar posicionar uma linha de regressão com os olhos não é apenas muito difícil, mas também inexato quando existem muitos pontos dispersos em uma área ampla. Felizmente, o coeficiente de regressão e o intercepto *y* podem ser obtidos matematicamente. O coeficiente de regressão (*b*) pode ser calculado a partir da covariância de *x* e *y* (cov_{xy}) e a variância de *x* (*s*) pela fórmula

$$b = \frac{cov_{xy}}{s_x^2} \tag{24.9}$$

O coeficiente de regressão indica o quanto *y* aumenta em média a cada aumento em *x*. A **Figura 24.14** ilustra várias linhas de

regressão com diferentes coeficientes de regressão. Observe que à medida que o coeficiente de regressão aumenta, a inclinação da linha de regressão aumenta.

Após calcular o coeficiente de regressão, o intercepto *y* pode ser calculado ao substituir o coeficiente de regressão e os valores médios de *x* e *y* na seguinte equação:

$$a = \bar{y} - b\bar{x} \tag{24.10}$$

A equação de regressão ($y = a + bx$, Equação 24.8) pode então ser usada para prever o valor de qualquer *y* considerando o valor de *x*.

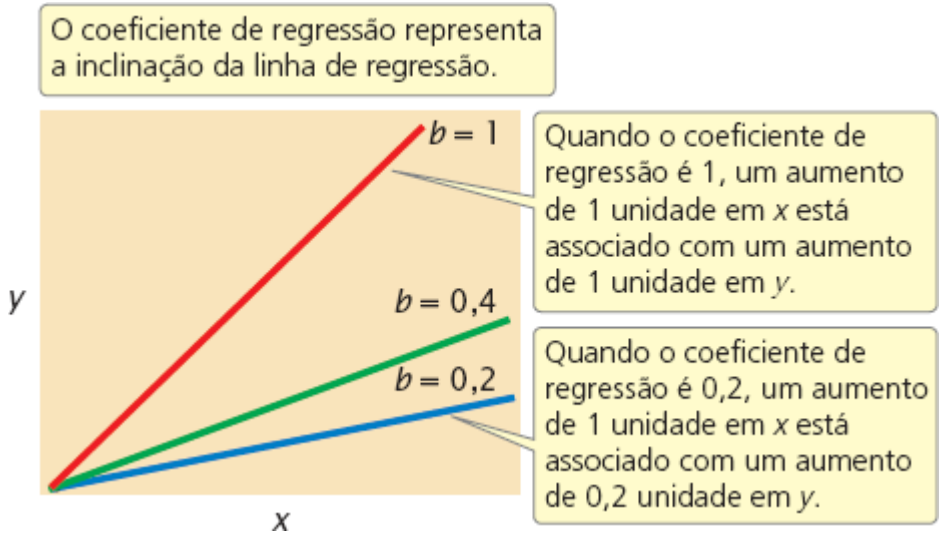


Figura 24.14 O coeficiente de regressão *b* representa a mudança em *y* por mudança de unidade em *x*. Aqui estão representadas as linhas de regressão com diferentes coeficientes de regressão.

Conceitos

Um coeficiente de correlação mede a força da associação entre duas variáveis. O sinal (positivo ou negativo) indica a direção da correlação, o valor absoluto mede a força da associação. A regressão é usada para prever o valor de uma variável com base no valor de uma variável correlacionada.

Checagem dos conceitos 4

Em Lubbock, Texas, a chuva e a temperatura exibem uma correlação significativa de –0,7. Qual conclusão está correta?

a. É comum ter mais chuva quando a temperatura está alta.
b. É comum ter mais chuva quando a temperatura está baixa.
c. A probabilidade de chover é a mesma quando a temperatura está alta ou baixa.

Problema desenvolvido

Os pesos corporais de 11 peixes fêmeas e o número de ovos que eles produzem são:

Peso (mg)	Ovos (milhares)
x	***y***
14	61
17	37
24	65
25	69
27	54
33	93
34	87
37	89

40	100
41	90
42	97

Quais são os coeficientes de correlação e de regressão para o peso corporal e número de ovos nestes 11 peixes?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

O coeficiente de correlação (r) e o coeficiente de regressão (b) para peso corporal e número de ovos nos 11 peixes.

Quais informações são fornecidas para solucionar o problema?

Pesos corporais e números de ovos para uma amostra de 11 peixes.

Etapas para a solução

Os cálculos necessários para responder a esta pergunta são fornecidos no quadro da próxima página. Para calcular os coeficientes de correlação e regressão, primeiro obtemos a soma de todos os valores x_i ($\Sigma\, x_i$) e a soma de todos os valores y_i ($\Sigma\, y_i$); estas somas são apresentadas na última coluna no quadro. Podemos calcular as médias de duas variáveis ao dividir as somas pelo número de medidas, que é 11:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{334}{11} = 30{,}36$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{842}{11} = 76{,}55$$

A **Peso (mg)**	**B**	**C**	**D** **Ovos (milhares)**	**E**	**F**	**G**
x	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
14	–16,36	267,65	61	–15,55	241,80	254,40
17	–13,36	178,49	37	–39,55	1.564,20	528,39
24	–6,36	40,45	65	–11,55	133,40	73,46
25	–5,36	28,73	69	–7,55	57,00	40,47
27	–3,36	11,29	54	–22,55	508,50	75,77
33	2,64	6,97	93	16,45	270,60	43,43
34	3,64	13,25	87	10,45	109,20	38,04
37	6,64	44,09	89	12,45	155,00	82,67
40	9,64	92,93	100	23,45	549,90	226,06
41	10,64	113,21	90	13,45	180,90	143,11
42	11,64	135,49	97	20,45	418,20	238,04
$\Sigma x_i = 334$		$\Sigma(x_i - \bar{x})^2 = 932{,}55$	$\Sigma y_i = 842$		$\Sigma(y_i - \bar{y})^2 = 4.188{,}70$	$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 1.743{,}84$

Fonte: R. R. Sokal and F. J. Rohlf, Biometry, 2 d ed. (San Francisco: W. H. Freeman and Company, 1981).

Após calcular as médias, os desvios de cada valor a partir da média são calculados, estes desvios estão apresentados nas

colunas B e E do quadro. Os desvios são então elevados ao quadrado (colunas C e F) e somados (última linha das colunas C e F). A seguir, os produtos do desvio dos valores *x* e o desvio dos valores *y* ($(x_i - \bar{x})\ (y_i - \bar{y})$) são calculados, estes produtos estão apresentados na coluna G e sua soma está apresentada na última linha da coluna G.

Para calcular a covariância, use a Equação 24.6:

$$\text{cov}_{xy} = \frac{\sum (x_i - \bar{x})\ (y_i - \bar{y})}{n - 1} = \frac{1.743,84}{10} = 174,38$$

Para calcular a correlação e a regressão são necessários as variâncias e desvios padrões de *x* e *y*:

$$s_x^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} = \frac{932,55}{10} = 93,26$$

$$s_x = \sqrt{s_x^2} = \sqrt{93,26} = 9,66$$

$$s_y^2 = \frac{\sum (y_i - \bar{y})^2}{n - 1} = \frac{4.188,70}{10} = 418,87$$

$$s_y = \sqrt{s_y^2} = \sqrt{418,87} = 20,47$$

Agora podemos calcular os coeficientes de correlação e regressão como apresentados aqui.

Coeficiente de correlação:

$$r = \frac{\text{cov}_{xy}}{s_x s_y} = \frac{174,38}{9,66 \times 20,47} = 0,88$$

Coeficiente de regressão:

$$b = \frac{\text{cov}_{xy}}{s_x^2} = \frac{174,38}{93,26} = 1,87$$

Agora teste sua compreensão sobre correlação e regressão ao resolver o **Problema 26** no final do capítulo.

Como aplicar a estatística no estudo de uma característica poligênica

Edward East fez um estudo estatístico inicial da herança poligênica sobre a altura das flores do tabaco (*Nicotiana longiflora*). Ele obteve duas variedades de tabaco com diferença na altura das flores: uma variedade tinha altura média de 40,5 mm e a outra tinha uma altura média da flor de 93,3 mm (**Figura 24.15**). Estas duas variedades foram intercruzadas por muitas gerações e eram homozigotas em todos os *loci* que contribuíam para a altura da flor. Assim, não havia variação genética nas cepas dos genitores originais; as pequenas diferenças na altura das flores dentro de cada cepa eram causadas por efeitos ambientais.

Quando East cruzou as duas cepas, ele descobriu que a altura da flor estava na metade entre as flores dos dois genitores (ver Figura 24.15), como seria esperado se os genes que determinam as diferenças nas duas cepas tivessem efeitos aditivos. A variância da altura da flor em F_1 foi semelhante à observada nos genitores, porque todo descendente F_1 tinha o mesmo genótipo, como tinha cada cepa parental (eram todos heterozigotos nos genes com diferenças entre as duas variedades parentais).

East cruzou F_1 e produziu os descendentes F_2. A altura média da flor de F_2 era semelhante à de F_1, mas a variância de F_2 era muito maior (ver Figura 24.15). Esta maior variabilidade indica que nem todos os descendentes de F_2 tinham o mesmo genótipo.

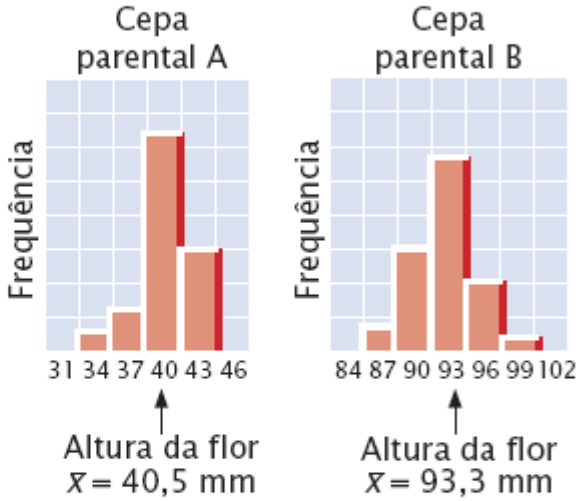
East selecionou algumas plantas de F_2 e as intercruzou para produzir os descendentes F_3. Ele descobriu que a altura da flor de F_3 dependia da altura da flor nas plantas selecionadas como suas genitoras. Este achado demonstrou que as diferenças de altura da flor em F_2 eram parcialmente genéticas e, portanto, foram transmitidas para a geração seguinte. Nenhuma das 444 plantas F_2 criadas por East exibiu alturas de flor semelhantes às das flores das duas cepas parentais. Este resultado sugeriu que mais de quatro *loci* com pares de alelos afetavam a altura das flores nas suas variedades, porque era esperado que quatro pares alélicos

produzissem 1 de 256 descendentes [$(1/4)^4 = 1/256$] tendo um dos fenótipos parentais originais.

Experimento

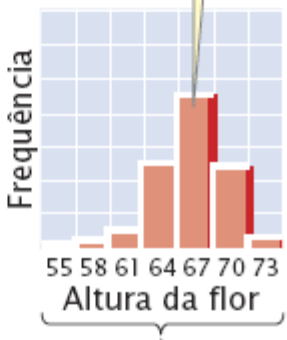
Pergunta: Como a altura da flor do tabaco é herdada?

Métodos **Geração P**



Resultados **Geração F_1**

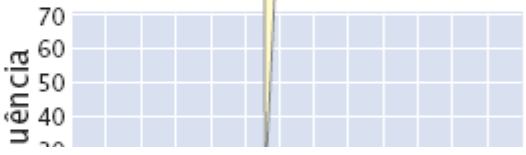
1 A altura da flor na F_1 estava entre a altura das flores dos genitores,...

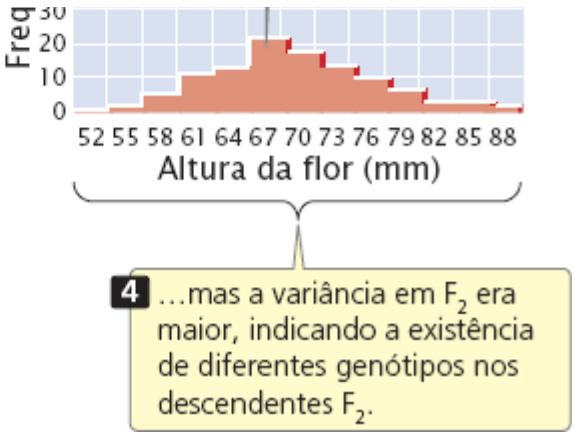


2 ...e a variância em F_1 era semelhante à observada nos genitores.

Geração F_2

3 A média de F_2 era semelhante à observada na F_1.





Conclusão: A altura da flor de F1 e F2 é consistente com a hipótese de que a característica é determinada por vários genes com efeitos aditivos.

Figura 24.15 Edward East fez um estudo estatístico inicial da herança da altura da flor no tabaco.

24.3 A herdabilidade é usada para estimar a proporção de variação em um traço genético

Além de serem poligênicas, as características quantitativas são frequentemente influenciadas por fatores ambientais. É útil saber o quanto da variação em uma característica quantitativa se deve a diferenças genéticas e o quanto se deve a diferenças ambientais. A proporção de variação fenotípica total devido a diferenças genéticas é conhecida como **herdabilidade**.

Considere um produtor de leite que tenha centenas de vacas leiteiras. Ele observa que algumas produzem mais leite que outras. A natureza destas diferenças é importante para o lucro de sua produção de leite. Se as diferenças na produção de leite têm, em sua maioria, origem genética, então o produtor pode ser capaz de aumentar a produção de leite ao reproduzir seletivamente as vacas que produzem mais leite. Por outro lado, se as diferenças têm, em sua maioria, origem ambiental, a reprodução seletiva terá pouco efeito na produção de leite e o produtor pode aumentar sua produção de leite ao ajustar os fatores ambientais associados com a maior produção de leite. Para determinar a extensão das influências genéticas e ambientais sobre a variação de uma característica, a variação fenotípica na característica deve ser separada em componentes atribuíveis a diferentes fatores.

Variância fenotípica

Para determinar o quanto de variação fenotípica em uma população se deve a fatores genéticos e o quanto se deve a fatores ambientais, primeiro precisamos ter alguma medida quantitativa do fenótipo em análise. Considere a população de plantas selvagens com diferentes tamanhos. Podemos coletar uma amostra representativa das plantas da população, pesar cada planta na amostra e calcular a média e variância do peso da planta. Esta **variância fenotípica** é representada por V_P.

Componentes da variância fenotípica. Primeiro, parte da variância fenotípica pode ocorrer por causa de diferenças nos genótipos entre membros individuais da população. Estas diferenças são chamadas de **variância genética** e são representadas por V_G.

Segundo, parte das diferenças no fenótipo pode ocorrer por causa de diferenças ambientais entre as plantas; essas diferenças são chamadas de **variância ambiental**, V_E. A variância ambiental inclui diferenças nos fatores ambientais como a quantidade de luz ou água que a planta recebe; também inclui diferenças aleatórias no desenvolvimento que não podem ser atribuídas a qualquer fator específico. Qualquer variação no fenótipo que não é herdado é, por definição, parte da variância ambiental.

Terceiro, a **variância da interação genética ambiental** (V_{GE}) surge quando o efeito de um gene depende do ambiente específico onde se encontra. A **Figura 24.16** apresenta um exemplo deste caso. Em um ambiente seco, o genótipo *AA* produz uma planta com média de 12 g de peso e o genótipo *aa* produz uma planta menor com média de 10 g. Em um ambiente úmido, o genótipo *aa* produz uma planta maior, variando 24 g de peso, enquanto o genótipo *AA* produz uma planta com média de 20 g. Neste exemplo, existem diferenças claras nos dois ambientes: ambos os genótipos produzem plantas mais pesadas no ambiente úmido. Também existem diferenças nos pesos de dois genótipos, mas os desempenhos relativos dos genótipos dependem de as plantas serem cultivadas em um ambiente seco ou úmido. Neste caso, as influências no fenótipo não podem ser organizadas em componentes genéticos e ambientais, porque a expressão do genótipo depende do ambiente onde a planta cresce. A variância fenotípica, portanto, deve incluir um componente que responda pela forma na qual os fatores genéticos e ambientais interagem.

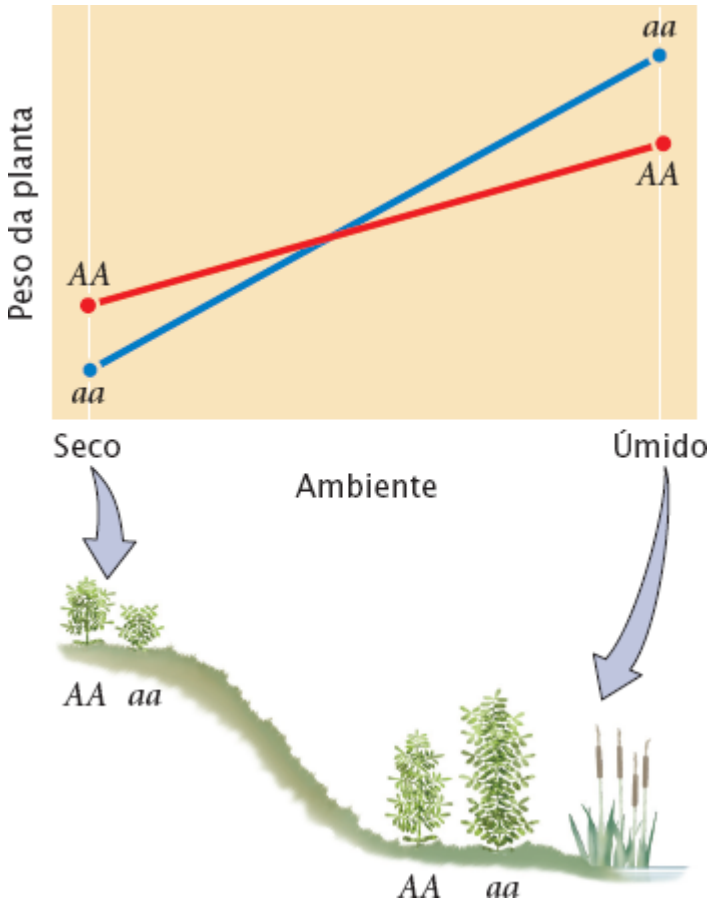


Figura 24.16 A variância da interação genética ambiental é obtida quando o efeito de um gene depende do ambiente específico onde se encontra. Neste exemplo teórico, o genótipo afeta o peso da planta, mas as condições ambientais determinam qual genótipo produz a planta mais pesada.

Em resumo, a variância fenotípica total pode ser distribuída em três componentes:

$$V_P = V_G + V_E + V_{GE} \tag{24.11}$$

Componentes da variância genética. A variância genética pode ser subdividida em componentes compostos de tipos diferentes de efeitos genéticos. Primeiro, a **variância genética aditiva** (V_A) inclui os efeitos aditivos dos genes sobre o fenótipo, que podem ser somados para determinar o efeito geral sobre o fenótipo. Por exemplo, suponha que, em uma planta, o alelo A^1 contribua com 2 g em peso e o alelo A^2 contribua 4 g. Se os alelos forem rigorosamente aditivos, então os genótipos teriam os seguintes pesos:

$$A^1A^1 = 2 + 2 = 4 \text{ g}$$
$$A^1A^2 = 2 + 4 = 6 \text{ g}$$
$$A^2A^2 = 4 + 4 = 8 \text{ g}$$

Os genes que Nilsson-Ehle estudou, que afetam a cor do grão do trigo, eram aditivos desta forma. A variância genética aditiva determina principalmente a semelhança entre os genitores e descendentes. Por exemplo, se toda a variância fenotípica é causada por variância genética aditiva, então o fenótipo médio dos descendentes será exatamente intermediário entre os dos genitores.

Segundo, existe **variância genética por dominância** (V_D) quando alguns genes têm um componente de dominância. Neste caso, os alelos em um *locus* não são aditivos; no caso, o efeito de um alelo depende da identidade do outro alelo naquele *locus*. Por exemplo, com um alelo dominante (*T*) os genótipos *TT* e *Tt* têm o mesmo fenótipo. Aqui, não podemos simplesmente adicionar os efeitos dos alelos, porque o efeito do alelo *t* para pequeno é mascarado pelo alelo *T* para grande. Em vez disso, é preciso adicionar um componente (V_D) à variância genética para se levar em conta no modo de interação dos alelos.

Terceiro, os genes em diferentes *loci* podem interagir da mesma forma que os alelos no mesmo *locus* interagem. Quando esta interação gênica ocorre, os efeitos dos genes não são aditivos. Por exemplo, o Capítulo 5 descreve como a cor do pelo do Labrador retriever exibe interação gênica, os genótipos *BB ee* e *bb ee* produzem cães com pelo amarelo, porque o efeito dos alelos no *locus B* são mascarados quando os alelos *ee* estão presentes no *locus E*. Com a interação gênica, é preciso adicionar

um terceiro componente, chamado **variância da interação gênica** (V_I), à variância genética:

$$V_G = V_A + V_D + V_I \tag{24.12}$$

Equação de resumo. Agora podemos integrar estes componentes em uma equação para representar todas as potenciais contribuições para a variância fenotípica:

$$V_P = V_A + V_D + V_I + V_E + V_{GE} \tag{24.13}$$

Esta equação nos fornece um modelo que descreve as potenciais causas de diferenças que observamos entre os fenótipos dos indivíduos. É importante observar que este modelo lida estritamente com *diferenças* observáveis (variância) nos fenótipos dos membros de uma população; ela não informa o valor absoluto da característica nem os genótipos envolvidos que produzem estas diferenças.

Tipos de herdabilidade

O modelo de variância fenotípica que acabamos de desenvolver pode ser usado para determinar o quanto da variância fenotípica em uma característica é causado por diferenças genéticas. A **herdabilidade de sentido amplo** (H^2) representa a proporção de variância fenotípica que é causada por variância genética e é calculada ao dividir a variância genética pela variância fenotípica:

$$\text{Herdabilidade no sentido amplo} = H^2 = \frac{V_G}{V_P} \tag{24.14}$$

O símbolo H^2 representa a herdabilidade de sentido amplo porque é uma medida da variância, que está em unidades ao quadrado.

A herdabilidade de sentido amplo pode variar de 0 a 1. Um valor de 0 indica que nenhuma variância fenotípica resulta das diferenças no genótipo e todas as diferenças no fenótipo resultam da variação ambiental. Um valor de 1 indica que toda variância fenotípica é resultado das diferenças nos genótipos. Um valor de herdabilidade entre 0 e 1 indica que ambos os fatores genéticos e ambientais influenciam a variância fenotípica.

Muitas vezes, estamos mais interessados na proporção de variância fenotípica que é o resultado da variância genética aditiva porque, como já foi mencionado, a variância genética aditiva determina principalmente a semelhança entre genitores e descendentes. A **herdabilidade de sentido restrito** (h^2) é igual à variância genética aditiva dividida pela variância fenotípica:

$$\text{Herdabilidade de sentido restrito} = h^2 = \frac{V_A}{V_P} \tag{24.15}$$

Resolva o Problema 27

Como calcular a herdabilidade

Considerando os componentes que contribuem para a variância fenotípica e desenvolvendo um conceito geral de herdabilidade, agora podemos perguntar como vamos estimar estes componentes diferentes e calcular a herdabilidade.

Existem várias formas para medir a herdabilidade de uma característica. Elas incluem eliminar um ou mais componentes de variância, comparar a semelhança dos genitores e dos descendentes, comparar as variâncias fenotípicas dos indivíduos com diferentes graus de relação e medir a resposta à seleção. A teoria matemática que envolve estes cálculos de herdabilidade é complexa, e então, aqui, vamos nos concentrar em desenvolver uma compreensão geral de como a herdabilidade é medida.

Herdabilidade pela eliminação dos componentes de variância. Uma forma de calcular a herdabilidade de sentido amplo consiste em eliminar um dos componentes da variância. Vimos que $V_P = V_G + V_E + V_{GE}$. Se eliminarmos toda variância ambiental ($V_E = 0$) então $V_{GE} = 0$ (porque se V_G ou V_E é zero, não pode ocorrer interação genética ambiental) e $V_P = V_G$. Em teoria, podemos fazer V_E igual a 0 ao garantir que todos os indivíduos foram criados no mesmo ambiente, mas na prática, isto é praticamente impossível. Por outro lado, se tornarmos V_G igual a 0 ao criar indivíduos geneticamente idênticos, V_P será igual a V_E. Em um típico experimento, podemos criar indivíduos clonados ou homozigotos idênticos em um ambiente definido e medir sua variância fenotípica para estimar V_E. Poderíamos criar um grupo de indivíduos geneticamente variáveis e medir sua variância fenotípica (V_P). Ao usar a V_E calculada nos indivíduos geneticamente idênticos, poderíamos obter a variância genética

dos indivíduos variáveis pela subtração:

$$V\text{G (de indivíduos geneticamente variáveis)} = V\text{P (de indivíduos geneticamente variáveis)} - V\text{E (de indivíduos geneticamente idênticos)} \quad (24.16)$$

A herdabilidade de sentido amplo de indivíduos geneticamente variáveis seria calculada assim:

$$H^2 = \frac{V_{\text{G (de indivíduos geneticamente variáveis)}}}{V_{\text{P (de indivíduos geneticamente variáveis)}}} \quad (24.17)$$

Sewall Wright usou este método para estimar a herdabilidade de manchas brancas em porquinhos-da-índia. Primeiro, ele mediu a variância fenotípica para a mancha branca em uma população geneticamente variável e descobriu que $V_P = 573$. Então ele cruzou os porquinhos por várias gerações de modo que eles fossem essencialmente homozigotos e idênticos do ponto de vista genético. Quando ele mediu sua variância fenotípica na mancha branca, ele obteve V_P igual a 340. Como $V_G = 0$ neste grupo, $V_P = V_G$. Wright assumiu este valor da variância ambiental para a população original (geneticamente variável) e estimou sua variância genética:

$$V_P - V_E = V_G$$
$$573 - 340 = 233$$

Ele estimou a herdabilidade de sentido amplo a partir das variâncias genética e fenotípica:

$$H^2 = \frac{V_G}{V_P}$$
$$H^2 = \frac{233}{573} = 0{,}41$$

Este valor indica que 41% da variação no padrão de mancha dos porquinhos-da-índia na população de Wright eram por causa de diferenças no genótipo.

A estimativa da herdabilidade com este método parte do pressuposto de que a variância ambiental de indivíduos geneticamente idênticos é a mesma que a variância ambiental de indivíduos geneticamente variáveis, o que pode não ser verdadeiro. Além disso, esta abordagem pode ser aplicada apenas em organismos em que é possível criar indivíduos geneticamente idênticos. Resolva o Problema 32

Herdabilidade pela regressão genitor-descendente

Outro método para estimar a herdabilidade é comparar os fenótipos de genitores e descendentes. Quando as diferenças genéticas são responsáveis pela variância fenotípica, os descendentes devem lembrar seus genitores mais do que eles lembram indivíduos distantes, porque os descendentes e genitores têm alguns genes em comum que ajudam a determinar seu fenótipo. A correlação e a regressão podem ser usadas para analisar a associação de fenótipos em diferentes indivíduos.

Para calcular a herdabilidade de sentido restrito desta forma, primeiro medimos a característica em uma série de genitores e descendentes. Os dados são organizados em famílias e o fenótipo parental médio é inserido contra o fenótipo médio dos descendentes (**Figura 24.17**). Cada ponto de dados no gráfico representa uma família, o valor no eixo *x* (horizontal) é o valor fenotípico médio dos genitores em uma família e o valor no eixo *y* (vertical) é o valor fenotípico médio dos descendentes da família.

Vamos considerar que não existe herdabilidade de sentido restrito para a característica ($h^2 = 0$), o significa que as diferenças genéticas não contribuem para as diferenças fenotípicas entre os indivíduos. Neste caso, os descendentes não serão mais semelhantes a seus genitores do que são com indivíduos distantes e os pontos estarão aleatoriamente dispersos, gerando um coeficiente de regressão igual a zero (ver Figura 24.17 A). A seguir, vamos considerar que todas as diferenças fenotípicas são causadas por diferenças genéticas aditivas ($h^2 = 1{,}0$). Neste caso, o fenótipo médio dos descendentes será igual ao fenótipo médio dos genitores e o coeficiente de regressão será 1 (ver Figura 24.17 B). Se os genes e o ambiente contribuírem para as diferenças no fenótipo, a herdabilidade e o coeficiente de regressão ficarão entre 0 e 1 (ver Figura 24.17 C). O coeficiente de regressão, portanto, fornece informações sobre a magnitude da herdabilidade.

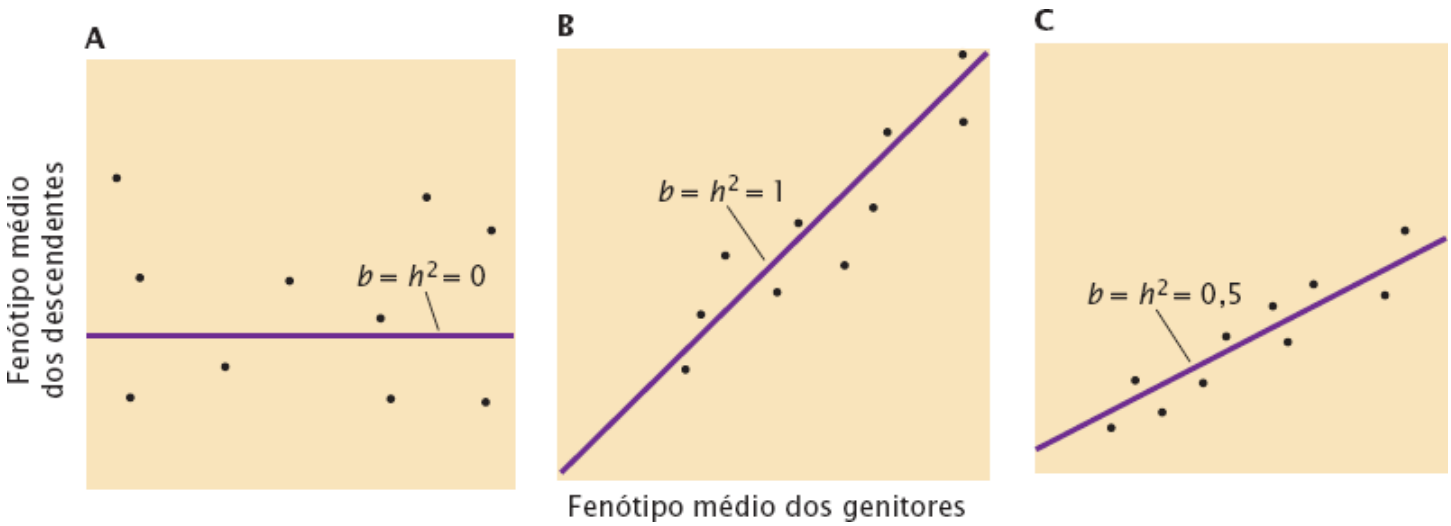


Figura 24.17 A herdabilidade de sentido restrito, h^2, é igual ao coeficiente de regressão, *b*, em uma regressão do fenótipo médio dos descendentes contra o fenótipo médio dos genitores. **A.** Não existe relação entre o fenótipo médio parental e o fenótipo médio dos descendentes. **B.** O fenótipo médio dos descendentes é o mesmo do fenótipo médio dos genitores. **C.** Ambos os genes e ambiente contribuem para as diferenças no fenótipo.

Uma complexa prova matemática (que não iremos apresentar aqui) demonstra que, em uma regressão do fenótipo médio dos descendentes contra o fenótipo médio dos genitores, a herdabilidade de sentido restrito (h^2) é igual ao coeficiente de regressão (b):

$$h^2 = b(\text{regressão da média dos descendentes contra a média de ambos os genitores}) \quad (24.18)$$

Um exemplo do cálculo da herdabilidade pela regressão dos fenótipos dos genitores e descendentes está ilustrado na **Figura 24.18**.

Em alguns casos, apenas o fenótipo de um genitor é conhecido. Em uma regressão do fenótipo médio dos descendentes contra o fenótipo de apenas um genitor, a herdabilidade de sentido restrito é igual a duas vezes o coeficiente de regressão:

$$h^2 = 2b(\text{regressão da média dos descendentes contra a média de um genitor}) \quad (24.19)$$

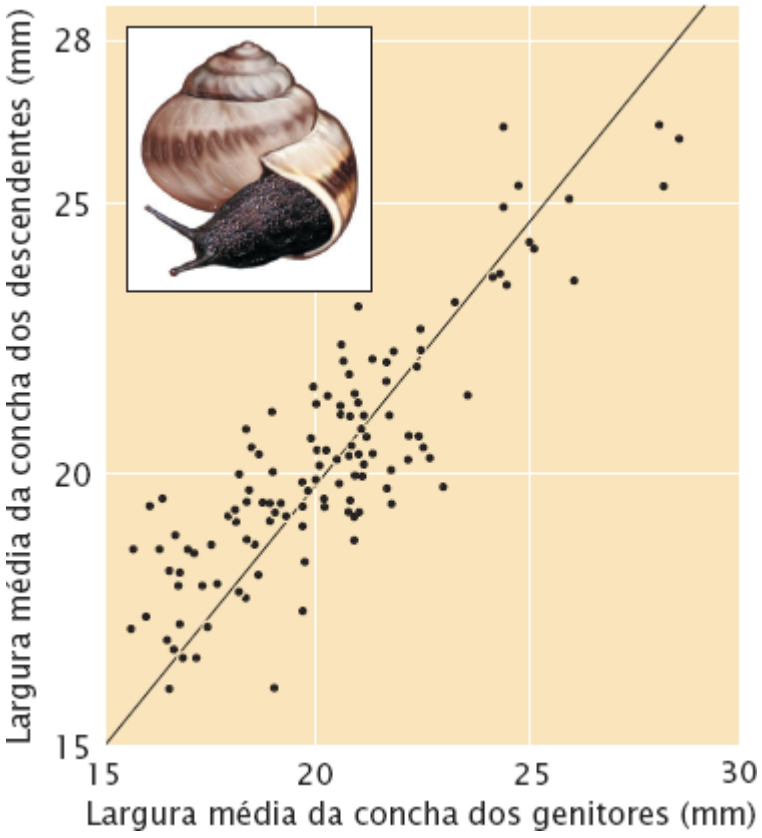


Figura 24.18 A herdabilidade da largura da concha nos caramujos pode ser determinada pela regressão do fenótipo médio dos descendentes contra o fenótipo médio dos genitores. O coeficiente de regressão, que é igual à herdabilidade, é 0,70. (De L.

M. Cook, Evolution 19:86-94, 1965.)

Com apenas um genitor, a herdabilidade é duas vezes o coeficiente de regressão porque apenas metade dos genes dos descendentes vem de um genitor, assim, é preciso dobrar o coeficiente de regressão para obter a herdabilidade completa.

Herdabilidade e graus de parentesco. Um terceiro método para calcular a herdabilidade é comparar os fenótipos dos indivíduos com diferentes graus de parentesco. Este método baseia-se no conceito de que quanto mais próximos são dois indivíduos, mais genes eles têm em comum.

Os gêmeos monozigóticos (idênticos) têm 100% dos seus genes em comum, enquanto os gêmeos dizigóticos (não idênticos) têm, em média, 50% dos seus genes em comum. Se os genes são importantes para determinar a variabilidade em uma característica, os gêmeos monozigóticos devem ser mais semelhantes em uma característica específica do que os gêmeos dizigóticos. Ao usar a correlação para comparar os fenótipos dos gêmeos monozigóticos e dizigóticos, podemos estimar a herdabilidade de sentido amplo. Uma estimativa aproximada da herdabilidade de sentido amplo pode ser obtida ao calcular o dobro da diferença dos coeficientes de correlação para uma característica quantitativa nos gêmeos monozigóticos e dizigóticos:

$$H^2 = 2(r_{MZ} - r_{DZ}) \tag{24.20}$$

em que r_{MZ} é igual ao coeficiente de correlação entre os gêmeos monozigóticos e r_{DZ} é igual ao coeficiente de correlação entre os gêmeos dizigóticos. Por exemplo, suponha que encontramos a correlação da altura entre os dois membros de pares de gêmeos monozigóticos (r_{MZ}) como sendo 0,9 e a correlação de altura entre os dois membros de gêmeos dizigóticos (r_{DZ}) como sendo 0,5. A herdabilidade de sentido amplo para a altura seria $H^2 = 2\ (0{,}9 - 0{,}5) = 2\ (0{,}4) = 0{,}8$. Este cálculo considera que os gêmeos monozigóticos crescem em ambientes que não são mais semelhantes do que o ambiente onde vivem dois indivíduos de um par de gêmeos dizigóticos. Esta suposição frequentemente não é satisfeita quando os gêmeos foram criados juntos.

A herdabilidade de sentido restrito também pode ser estimada ao comparar as variâncias fenotípicas para uma característica em dois irmãos consanguíneos (mesmos genitores em comum assim como uma média de 50% de seus genes) e meios-irmãos (que têm um genitor em comum e uma média de 25% de seus genes).

Todas as estimativas de herdabilidade dependem da suposição de que os ambientes dos indivíduos com grau de parentesco não são semelhantes aos dos indivíduos que não são próximos. É difícil cumprir esta suposição em estudos com humanos, porque em geral as pessoas com grau de parentesco crescem juntas. As estimativas de herdabilidade para humanos devem sempre ser observadas com cuidado. Resolva o Problema 36

Conceitos

A herdabilidade de sentido amplo é a proporção de variância fenotípica devida à variância genética. A herdabilidade de sentido restrito é a proporção de variância fenotípica devida à variância genética aditiva. A herdabilidade pode ser medida ao eliminar um dos componentes da variância, ao analisar a regressão genitor-descendente ou comparar indivíduos com diferentes graus de parentesco.

Checagem dos conceitos 5

Se a variância ambiental (V_E) aumentar e todos os outros componentes de variância permanecerem iguais, qual seria o efeito?

a. A herdabilidade de sentido amplo reduzirá.
b. A herdabilidade de sentido amplo aumentará.
c. A herdabilidade de sentido restrito aumentará.
d. A herdabilidade de sentido amplo aumentará, mas a herdabilidade de sentido restrito reduzirá.

Limitações da herdabilidade

Conhecer a herdabilidade é muito prático porque permite que possamos prever estatisticamente os fenótipos dos descendentes com base no fenótipo de seus genitores. Também informa como as características responderão à seleção (ver Seção 24.4). Apesar da sua importância, a herdabilidade é frequentemente mal compreendida. Ela não fornece informações sobre os genes de um indivíduo ou dos fatores ambientais que controlam o desenvolvimento de uma característica nem sobre a natureza das diferenças entre os grupos. Esta seção destaca algumas limitações e ideias erradas comuns sobre a herdabilidade de sentido amplo e restrito.

A herdabilidade não indica o quanto uma característica é geneticamente determinada. A herdabilidade é a

proporção de variância fenotípica devido à variância genética; ela não diz nada sobre o grau no qual os genes determinam uma característica. A herdabilidade indica apenas o quanto os genes determinam a *variação* em uma característica. A determinação de uma característica e a determinação da variação em uma característica são duas coisas muito diferentes.

Considere a polidactilia (existência de dedos extras) nos coelhos, que pode ser causada por fatores ambientais ou um gene dominante. Suponha que temos um grupo de coelhos todos homozigotos para um gene que produz o número comum de dedos. Nenhum dos coelhos neste grupo possui um gene para polidactilia, mas alguns dos coelhos são polidactilos por causa de fatores ambientais. A herdabilidade de sentido amplo para a polidactilia neste grupo é zero, porque não existe variação genética para polidactilia; toda variação deve-se a fatores ambientais. Entretanto, seria incorreto concluirmos que os genes não participam na determinação no número de dígitos nos coelhos. De fato, sabemos que existem alelos específicos que podem produzir dedos extras (embora estes alelos não existam no grupo de coelhos em análise). A herdabilidade não indica nada sobre os genes que controlam o desenvolvimento de uma característica; ela fornece informações apenas sobre casos de variação de uma característica em um grupo definido.

Um indivíduo não tem herdabilidade. As herdabilidades de sentido amplo e restrito são valores estatísticos com base em variâncias genéticas e fenotípicas encontradas em um *grupo* de indivíduos. A herdabilidade não pode ser calculada para um indivíduo e ela não tem significado para um indivíduo específico. Suponha que calculamos a herdabilidade de sentido restrito de peso corporal adulto de estudantes em uma turma de biologia e obtemos o valor 0,6. Podemos concluir que 60% da variação no peso corporal adulto entre os estudantes nesta classe são determinados pela variação genética aditiva. Não devemos, entretanto, concluir que 60% do peso corporal de qualquer estudante se devam a genes aditivos.

Não há herdabilidade universal para uma característica. O valor da herdabilidade para uma característica é específica para uma dada população em um determinado ambiente. Lembre-se de que a herdabilidade de sentido amplo é a variância genética dividida pela variância fenotípica. A variância genética depende de quais alelos estão presentes, que são diferentes entre as populações. No exemplo da polidactilia nos coelhos, não existiam alelos para polidactilia no grupo, então a herdabilidade da característica era zero. Um grupo diferente de coelhos pode ter muitos alelos para polidactilia e a herdabilidade da característica pode ser elevada.

As diferenças ambientais também podem afetar a herdabilidade porque V_P é composta por variância genética e ambiental. Quando as diferenças ambientais que afetam uma característica são diferentes entre dois grupos, as herdabilidades para os dois grupos também podem ser diferentes.

Como a herdabilidade é específica para uma população definida em um dado ambiente, é importante não extrapolar esse coeficiente de uma população para outra. Por exemplo, a altura do humano é determinada por fatores ambientais (como nutrição e saúde) e pelos genes. Se medirmos a herdabilidade da altura em um país desenvolvido, podemos obter um valor 0,8, indicando que a variação na altura nesta população é, em grande parte, genética. Esta população tem uma alta herdabilidade porque a maioria das pessoas tem nutrição e cuidados de saúde adequados (V_E é baixa), então a maior parte da variação fenotípica na altura é determinada pela genética. Seria incorreto considerarmos que a altura tem uma alta herdabilidade em todas as populações humanas. Nos países em desenvolvimento, pode ter mais variação em uma faixa de fatores ambientais, algumas pessoas podem ter boa nutrição e saúde, enquanto outras podem ter uma dieta deficiente em proteínas e sofrer de doenças que afetam a altura. Se medirmos a herdabilidade de altura nesse país, sem dúvida teríamos um valor menor do que observamos no país desenvolvido porque existe mais variação ambiental e a variância genética em altura é uma proporção menor de variação fenotípica, reduzindo a herdabilidade. O ponto importante é lembrar que a herdabilidade tem de ser calculada separadamente para cada população e cada ambiente.

Mesmo quando a herdabilidade é alta, os fatores ambientais podem influenciar uma característica. Alta herdabilidade não significa que os fatores ambientais não possam influenciar a expressão de uma característica. A alta herdabilidade indica apenas que a variação ambiental à qual a população está *atualmente* exposta não é responsável pela variação na característica. Vamos examinar de novo a altura nos humanos. Na maioria dos países desenvolvidos, a herdabilidade da altura é alta, indicando que as diferenças genéticas são responsáveis pela maior parte da variação da altura. Seria errado concluirmos que a altura nos humanos não pode ser mudada pela alteração do ambiente. Na verdade, a altura diminuiu em várias cidades europeias durante a Segunda Guerra Mundial por causa da fome e doença e pode ser aumentada drasticamente pela administração de hormônio do crescimento (GH) nas crianças. A ausência de variação ambiental em uma característica não significa que a característica não responderá a uma mudança ambiental.

A herdabilidade não revela nada sobre a natureza das diferenças de população de uma característica. Um conceito errado comum sobre a herdabilidade é que ela fornece informações sobre as diferenças de uma característica em uma população. A herdabilidade é específica para uma dada população em determinado ambiente, então não pode ser usada para tirar

conclusões sobre como as populações diferem em uma característica.

Suponha que determinemos a herdabilidade para a altura em humanos em dois grupos. Um grupo é de uma pequena cidade em um país desenvolvido, onde todos consomem uma dieta rica em proteínas. Como existe pouca variação nos fatores ambientais que afetam a altura e existe variação genética, a herdabilidade da altura neste grupo é alta. O segundo grupo inclui os habitantes de uma vila simples em um país em desenvolvimento. O consumo de proteína por estas pessoas é apenas 25% do que é consumido pelo primeiro grupo, então a estatura média dos adultos é vários centímetros a menos do que no país desenvolvido. Novamente, existe pouca variação nos fatores ambientais que determinam a altura neste grupo, porque todos na vila comem os mesmos tipos de alimentos e estão expostos às mesmas doenças. Como existe pouca variação ambiental e existe alguma variação genética, a herdabilidade da altura neste grupo também é alta.

Assim, a herdabilidade em ambos os grupos é alta, e a altura média nos dois grupos é consideravelmente diferente. Podemos ficar tentados a concluir que a diferença na altura entre os dois grupos depende da genética, que as pessoas do país desenvolvido são geneticamente maiores do que as pessoas do país em desenvolvimento. É claro que esta conclusão está errada, pois estas diferenças na altura são devidas a dieta, um fator ambiental. A herdabilidade não fornece informações sobre as causas das diferenças entre as populações.

Estas limitações da herdabilidade são ignoradas, particularmente em argumentos sobre as possíveis implicações sociais de diferenças genéticas entre os humanos. Logo após os princípios de Mendel serem redescobertos, alguns geneticistas começaram a alegar que muitas características comportamentais são determinadas inteiramente pelos genes. Esta afirmação levou a debates sobre se as características como a inteligência humana são determinadas por genes ou pelo ambiente. Muitas das afirmações iniciais do comportamento humano oriundo da genética baseavam-se em pesquisa insatisfatória. Infelizmente, os resultados destes estudos foram aceitos na íntegra e levaram a várias leis eugênicas que discriminaram contra alguns grupos de pessoas. Atualmente, os geneticistas reconhecem que muitas características comportamentais são influenciadas por uma interação complexa de genes e ambiente, e é muito difícil separar os efeitos genéticos dos efeitos do ambiente.

Os resultados de vários estudos modernos indicam que a inteligência humana medida com QI e outros testes de inteligência tem uma herdabilidade moderadamente alta (em geral de 0,4 a 0,8). Com base nesta observação, algumas pessoas argumentam que a inteligência é inata e que aumentar as oportunidades de educação não podem acentuar a inteligência. Este argumento baseia-se na ideia errada de que, quando a herdabilidade é alta, mudar o ambiente não modificará a característica. Além disso, como a herdabilidade de inteligência varia entre 0,4 e 0,8, uma parte considerável da variância na inteligência é consequente a diferenças ambientais.

Outro argumento com base na ideia errada sobre a herdabilidade é que as diferenças étnicas nas medidas de inteligência estão relacionadas com a genética. Como os resultados de alguns estudos genéticos mostram que o QI tem herdabilidade moderadamente alta e outros estudos encontraram diferenças no QI médio de grupos étnicos, algumas pessoas sugeriram que as diferenças étnicas no QI são genéticas. Como no exemplo dos efeitos da dieta na altura dos humanos, a herdabilidade não fornece informações sobre as causas de diferenças entre os grupos; ela indica apenas o grau no qual a variância fenotípica em um único grupo baseia-se na genética. A alta herdabilidade para uma característica não significa que as diferenças fenotípicas entre os grupos étnicos sejam genéticas. Também devemos lembrar que é muito difícil separar os efeitos genéticos e ambientais nos humanos, então a herdabilidade estimada por si só pode não ser confiável. Resolva o Problema 35

Conceitos

A herdabilidade fornece informações apenas sobre o grau no qual a variação em uma característica é determinada pela genética. Não existe herdabilidade universal para uma característica; ela é específica para uma dada população em um dado ambiente. Os fatores ambientais podem afetar potencialmente as características com alta herdabilidade e a herdabilidade não diz nada sobre a natureza das diferenças de população em uma característica.

Checagem dos conceitos 6

Suponha que você acabou de aprender que a herdabilidade de sentido restrito da pressão arterial medida em um grupo de afro-americanos em Detroit, Michigan, é 0,4. O que a herdabilidade pode nos dizer sobre as contribuições genéticas e ambientais para a pressão arterial?

Como localizar os genes que afetam as características quantitativas

Os métodos estatísticos descritos para uso na análise de características quantitativas podem ser usados para fazer previsões sobre o fenótipo médio esperado nos descendentes e estimar a contribuição geral dos genes para a variação na característica. Estes métodos, entretanto, não permitem que identifiquemos e determinemos a influência dos genes individuais que afetam as

características quantitativas. Como foi discutido no início deste capítulo, as regiões cromossômicas com genes que controlam as características poligênicas são chamadas de *loci* de traço quantitativo. Embora a genética quantitativa tenha feito importantes contribuições para a biologia básica e para a reprodução de plantas e animais, a antiga incapacidade de identificar QTLs e medir seus efeitos individuais limitou em muito a aplicação de métodos genéticos quantitativos.

Como mapear QTLs. Recentemente, foram identificados e mapeados numerosos marcadores genéticos com o uso de técnicas moleculares, possibilitando identificar QTLs com análise de ligação. A ideia em questão é simples: se a herança de um marcador genético está associada consistentemente com a herança de uma característica particular (como a altura maior), então o marcador deve estar ligado a um QTL que afeta a altura. A chave é ter marcadores genéticos suficientes de modo que os QTLs possam ser detectados por todo genoma. Com a introdução de polimorfismos de comprimento do fragmento de restrição, variações de microssatélites e polimorfismos de nucleotídio único (SNPs, ver os *Capítulos 19* e *20*), marcadores variáveis estão disponíveis para mapear QTLs em vários organismos diferentes (**Figura 24.19**).

Um procedimento comum para mapear QTLs é cruzar duas cepas homozigotas com diferentes alelos em muitos *loci* (**Figura 24.20**). Os descendentes F_1 resultantes são intercruzados ou retrocruzados para permitir que os genes recombinem-se por meio da segregação independente e *crossing over*. Os genes em diferentes cromossomos e os genes distantes no mesmo cromossomo vão se recombinar livremente, os genes que estão intimamente ligados vão ser herdados juntos. Os descendentes são medidos para uma ou mais características quantitativas, ao mesmo tempo, é feito seu genótipo para pesquisar numerosos marcadores genéticos que se estendem pelo genoma. Qualquer correlação entre a herança de um alelo marcador específico e um fenótipo quantitativo indica que um QTL está ligado àquele marcador. Se forem usados marcadores suficientes, teoricamente é possível detectar todos os QTLs que afetam uma característica. É importante reconhecer que um QTL não é um gene; ao contrário, é uma localização do mapa para uma região cromossômica associada com aquele traço. Após um QTL ser identificado, ele pode ser estudado quanto à existência de um ou mais genes específicos ou outras sequências que influenciam o traço quantitativo. A introdução deste capítulo descreve como esta abordagem foi usada para identificar um gene principal que afeta a produção de óleo no milho. O mapeamento do QTL é usado para detectar genes que afetam várias características nas espécies de plantas e animais (**Quadro 24.2**).

Figura 24.19 O mapeamento de QTLs é usado para identificar genes que influenciam muitos traços quantitativos importantes, incluindo a massa muscular nos porcos. (USDA.)

Estudos de associação genômica ampla. O método tradicional para identificar QTLs é feito para promover cruzamentos entre variedades com diferentes traços, e então faz o genótipo de numerosos descendentes para muitos marcadores. Embora seja eficiente, este método é lento e trabalhoso.

Uma técnica alternativa de identificação dos genes que afetam os traços quantitativos consiste em fazer estudos de associação genômica ampla, que foram apresentados no Capítulo 7. Diferente da análise de ligação tradicional, que examina a associação de um traço e marcadores de genes entre os *descendentes de um cruzamento*, os estudos de associação genômica ampla procuram por associações entre traços e marcadores genéticos em uma *população biológica*, um grupo de indivíduos cruzados. A presença de uma associação entre os marcadores genéticos e um traço indica que os marcadores genéticos estão intimamente

ligados a um ou mais genes que afetam a variação no traço. Os estudos de associação genômica ampla foram facilitados pela identificação de polimorfismos de nucleotídio único, que são posições no genoma onde os organismos individuais variam em um único par de bases (ver *Capítulo 20*). Foram identificados muitos SNPs em organismos por meio do sequenciamento genômico. Organismos individuais podem ser genotipados de forma rápida e barata para numerosos SNPs, fornecendo os marcadores genéticos necessários para fazer estudos de associação genômica ampla.

Quadro 24.2 Exemplos de características quantitativas para as quais os QTLs foram detectados.

Organismo	Característica quantitativa
Tomate	Sólidos solúveis Peso da fruta pH da fruta Crescimento Formato da folha Altura
Milho	Altura Comprimento da folha Número de perfilhares Firmeza da gluma Rendimento de grãos Número de orelhas Termotolerância
Feijão-comum	Número de nódulos
Feijão-da-china	Peso da semente
Feijão-de-corda	Peso da semente
Trigo	Broto pré-colheita
Porco	Crescimento Comprimento do intestino delgado Gordura média do dorso Gordura abdominal
Camundongo	Epilepsia
Rato	Hipertensão arterial

Fonte: De S. D. Tanksley, Mapping polygenes, Annual Review of Genetics 27:218, 1993.

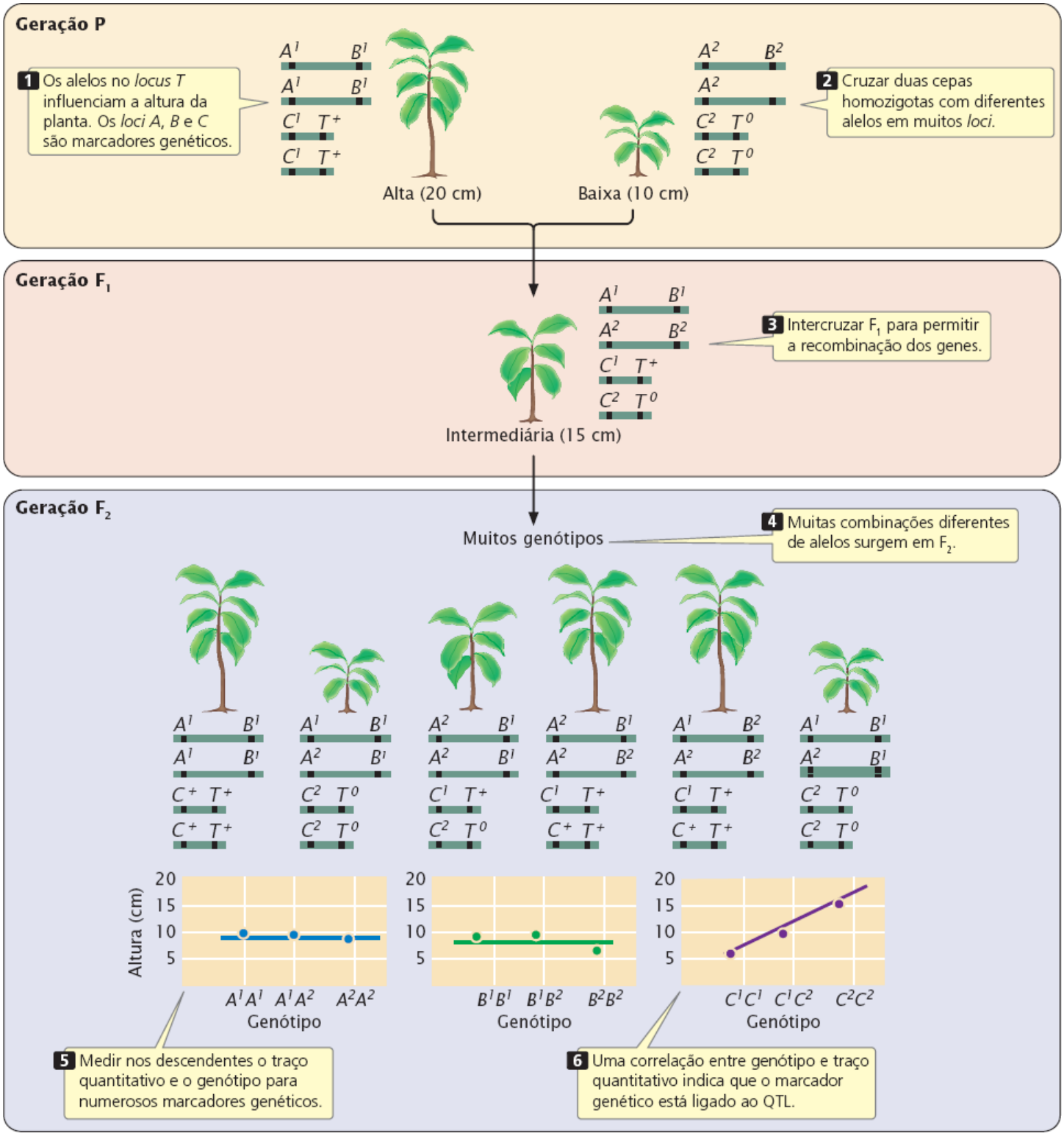


Figura 24.20 O mapeamento dos *loci* de traço quantitativo pela análise de ligação pode ajudar a identificar os genes que ajudam a determinar as diferenças nos traços quantitativos. Os genótipos no *locus C* estão associados com a herança de diferenças na altura da planta, indicando que um QTL para altura (o *locus T*) está intimamente ligado ao *locus C*.

Estes estudos são muito usados para localizar genes que afetam os traços quantitativos nos humanos, incluindo a suscetibilidade a doença, obesidade, inteligência e altura. Também foram estudados vários traços quantitativos, incluindo composição, tamanho, cor e gosto do grão, resistência a doenças e qualidade do amido. Os estudos de associação genômica ampla em animais domésticos identificaram segmentos cromossômicos que afetam o peso corporal, a composição corporal, os traços reprodutivos, os níveis de hormônio, as características dos pelos e o comportamento.

Conceitos

A disponibilidade de numerosos marcadores genéticos revelada por métodos moleculares possibilitou mapear segmentos cromossômicos com genes que contribuem para características poligênicas. Os estudos de associação genômica ampla localizam genes que afetam traços quantitativos ao detectar associações

entre os marcadores genéticos e um traço em uma população.

24.4 Os traços geneticamente variáveis mudam em resposta à seleção

A evolução é a mudança genética que ocorre nos membros de uma população com o passar do tempo. Várias forças diferentes são potencialmente capazes de levar à evolução e vamos explorar estas forças e o processo de evolução com mais detalhes no Capítulo 25. Aqui, consideramos como uma destas forças, a seleção natural, pode produzir a mudança genética em uma característica quantitativa.

Charles Darwin propôs a ideia de seleção natural em seu livro *A Origem das Espécies* em 1859. A **seleção natural** surge pela reprodução diferencial de indivíduos com diferentes genótipos. Graças a seus genes, alguns indivíduos produzem mais descendentes que outros. Os reprodutores com mais sucesso dão origem a mais descendentes, que herdam os genes que conferem uma vantagem reprodutiva. Assim, as frequências dos genes que conferem uma vantagem reprodutiva aumentam com o tempo e a população evolui. A seleção natural é uma das forças mais importantes que produzem a mudança evolutiva. Pela seleção natural, os organismos se tornam geneticamente adaptados a seus ambientes; à medida que os ambientes mudam, os grupos de organismos mudam as formas nas quais eles são mais bem adaptados a sobreviver e reproduzir.

Por milhares de anos, os humanos praticaram uma forma de seleção ao promover a reprodução de organismos com traços desejáveis. Esta forma de seleção é a **seleção artificial** e produziu as plantas e animais domésticos que possibilitaram a agricultura moderna. A força da seleção artificial, a primeira aplicação dos princípios genéticos pelos humanos, está ilustrada pela incrível diversidade de formas, cores e comportamentos dos cães domesticados modernos (**Figura 24.21**).

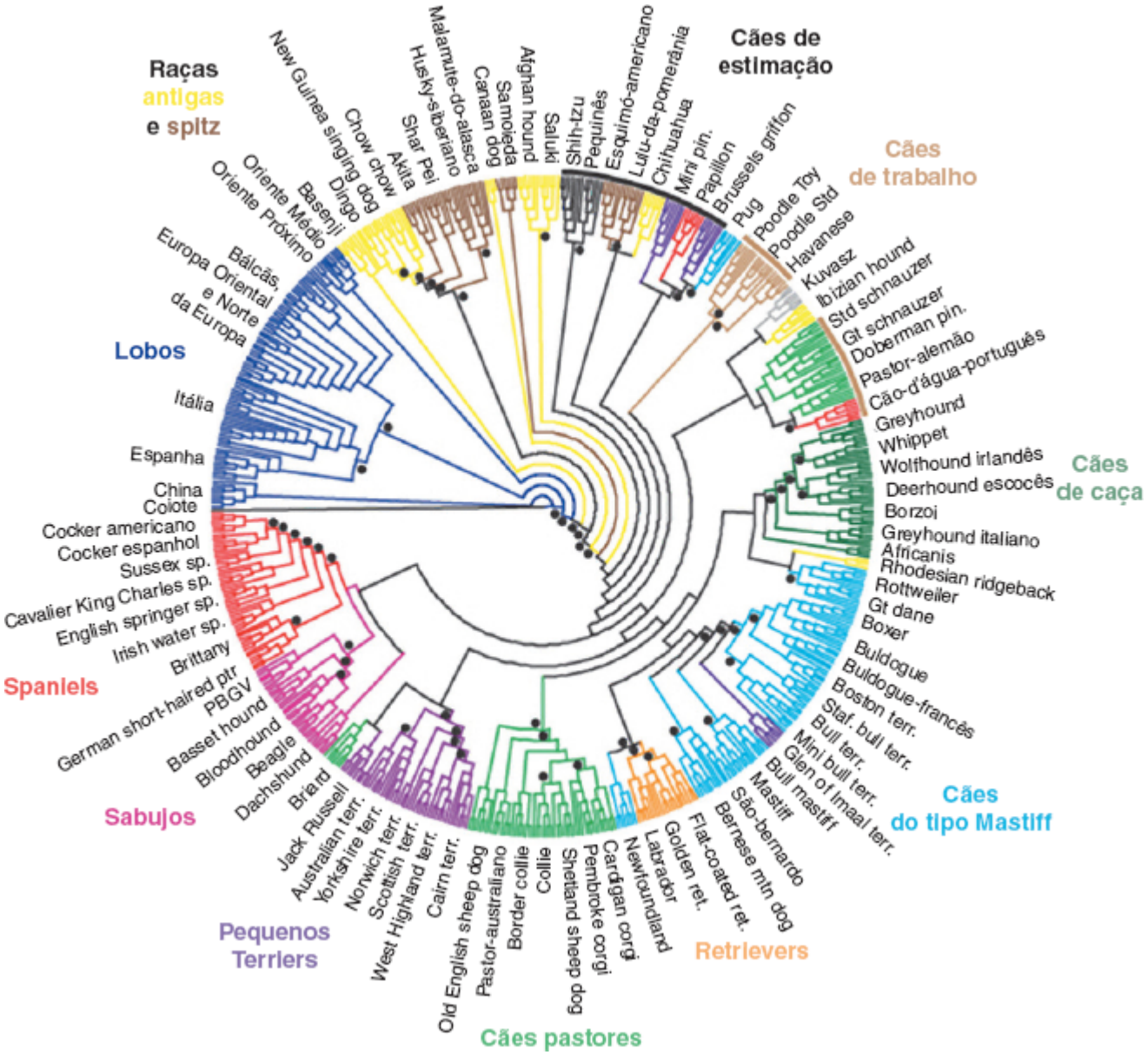


Figura 24.21 A seleção artificial produziu a incrível diversidade de forma, tamanho, cor e comportamento observados hoje em dia nas raças de cães domésticos. Este diagrama descreve as relações evolutivas entre os lobos e as diferentes raças de cães a

partir das análises de sequências de DNA. (Reimpressa com permissão de Macmillan Publishers Ltd.: B.M. von Holdt et al., Nature, 464 [7290], 898-902. © 2010.)

Como prever a resposta à seleção

Quando uma característica quantitativa está sujeita a seleção natural ou artificial, frequentemente ela muda com o tempo, desde que exista variação genética para a característica na população. Suponha que um fazendeiro queira aumentar a produção de leite entre as vacas do seu rebanho. A variação em vários *loci* afeta potencialmente a produção de leite nas vacas; alguns alelos nestes *loci* conferem produção elevada de leite, enquanto outros alelos conferem baixa produção. O produtor cruza apenas as vacas que produzem mais leite. Se existir variação genética na produção de leite (ou seja, existem diferentes alelos nos *loci* que controlam a produção de leite), a produção média de leite nos descendentes das vacas selecionadas será maior que a produção média do rebanho original. Esta maior produção se deve ao fato de que as vacas selecionadas têm mais alelos para a produção elevada de leite do que a média das vacas e estes alelos são transmitidos para os descendentes. Os descendentes das vacas selecionadas têm maior proporção de alelos para maior rendimento de leite e, portanto, produzem mais leite que a média das vacas do rebanho original.

O quanto uma característica sujeita a seleção muda em uma geração é chamada de **resposta à seleção**. Suponha que a vaca média em um rebanho leiteiro produza 80 ℓ de leite por semana. Um fazendeiro seleciona, para aumentar a produção de leite, por reprodução, as maiores produtoras de leite, e as descendentes destas vacas selecionadas produzem 100 ℓ de leite por semana em média. A resposta à seleção é calculada ao subtrair o fenótipo médio da população original (80 ℓ) do fenótipo médio dos descendentes (100 ℓ), obtendo uma seleção de 100 – 80 = 20 ℓ por semana.

Fatores que influenciam a resposta à seleção. A resposta à seleção é determinada principalmente por dois fatores. Primeiro, é afetada pela herdabilidade de sentido restrito, que determina em grande parte o grau de semelhança entre os genitores e descendentes. Quando a herdabilidade de sentido restrito é alta, os descendentes tendem a lembrar seus genitores, por outro lado, quando esta herdabilidade é baixa, existe pouca semelhança entre genitores e descendentes.

O segundo fator que determina a resposta à seleção é o quanto há de seleção. Se o fazendeiro for muito rigoroso na escolha dos genitores e cruzar apenas as maiores produtoras de leite no rebanho (ou seja, as duas principais vacas), todos os descendentes receberão os genes para a produção de leite de alta qualidade. Se o fazendeiro for menos seletivo e cruzar as 20 principais produtoras no rebanho, então os descendentes não carrearão tantos genes superiores para a alta produção de leite e em média não produzirão tanto leite como os descendentes das duas principais produtoras. A resposta à seleção depende da diferença fenotípica dos indivíduos que são selecionados como genitores; esta diferença fenotípica é medida pela **seleção diferencial**, definida como a diferença entre o fenótipo médio dos genitores selecionados e o fenótipo médio da população original. Se a produção média de leite do gado original for de 80 ℓ e o fazendeiro cruzar as vacas com uma produção média de leite de 120 ℓ, então a seleção diferencial será 120 – 80 = 40 ℓ.

Cálculo da resposta à seleção. A resposta à seleção (R) depende da herdabilidade de sentido restrito (h^2) e da seleção diferencial (S):

$$R = h^2 \times S \tag{24.21}$$

Esta equação pode ser usada para prever a magnitude de mudança em uma característica quando é aplicada a seleção diferencial. G. A. Clayton e colaboradores estimaram a resposta à seleção que ocorreria no número de pelos abdominais da *Drosophila melanogaster*. Ao usar vários métodos diferentes, incluindo a regressão genitor-descendente, eles primeiro estimaram a herdabilidade de sentido restrito do número de cerdas abdominais em uma população de moscas-da-fruta como sendo 0,52. O número médio de cerdas na população original era 35,3. Eles selecionaram moscas com um número médio de cerda de 40,6 e as intercruzaram para produzir a geração seguinte. A seleção diferencial era 40,6 – 35,3 = 5,3; então eles previram uma resposta à seleção como sendo

$$R = 0{,}52 \times 5{,}3 = 2{,}8$$

A resposta à seleção de 2,8 é o aumento esperado na característica dos descendentes acima da média da população original. Eles esperavam um número médio de cerdas abdominais nos descendentes das suas moscas selecionadas como sendo 35,3 + 2,8 = 38,1. De fato, eles encontraram um número médio de cerdas de 37,9 nestas moscas.

Como estimar a herdabilidade a partir da resposta à seleção. Ao reorganizar a Equação 24.21, teremos outra forma para calcular a herdabilidade de sentido restrito:

$$h^2 = \frac{R}{S} \quad (24.22)$$

Desta forma, h^2 pode ser calculada ao conduzir um experimento de resposta à seleção. Primeiro a seleção diferencial é obtida ao subtrair a média da população da média dos genitores selecionados. Os genitores selecionados são então cruzados e o fenótipo médio de seus descendentes é calculado. A diferença entre a média dos descendentes e da população inicial é a resposta à seleção, que pode ser usada com a seleção diferencial para estimar a herdabilidade. A herdabilidade determinada por um experimento de resposta à seleção é chamada de **herdabilidade realizada**. Se algumas suposições são satisfeitas, a herdabilidade realizada é idêntica à herdabilidade de sentido restrito.

Um dos experimentos de seleção mais longos é um estudo de teor de óleo e proteína nas sementes do milho (**Figura 24.22**). Este experimento começou na University of Illinois em 163 espigas de milho com um teor de óleo variando de 4 a 6%. As plantas de milho com teor elevado de óleo e com teor menor foram selecionadas e cruzadas. A resposta à seleção para o maior teor de óleo (a linha superior da Figura 24.22) alcançou cerca de 20%, enquanto a resposta à seleção para o menor teor de óleo alcançou um limite inferior próximo de zero. As análises genéticas de cepas com alto e baixo teores de óleo revelaram que pelo menos 20 *loci* participam da determinação do teor de óleo, um dos quais exploramos na introdução deste capítulo. Resolva o Problema 40

Conceitos

A resposta à seleção é influenciada pela herdabilidade de sentido restrito e pela seleção diferencial.

Checagem dos conceitos 7

A herdabilidade de sentido restrito para um traço é 0,4 e a seleção diferencial é 0,5. Qual é a resposta prevista à seleção?

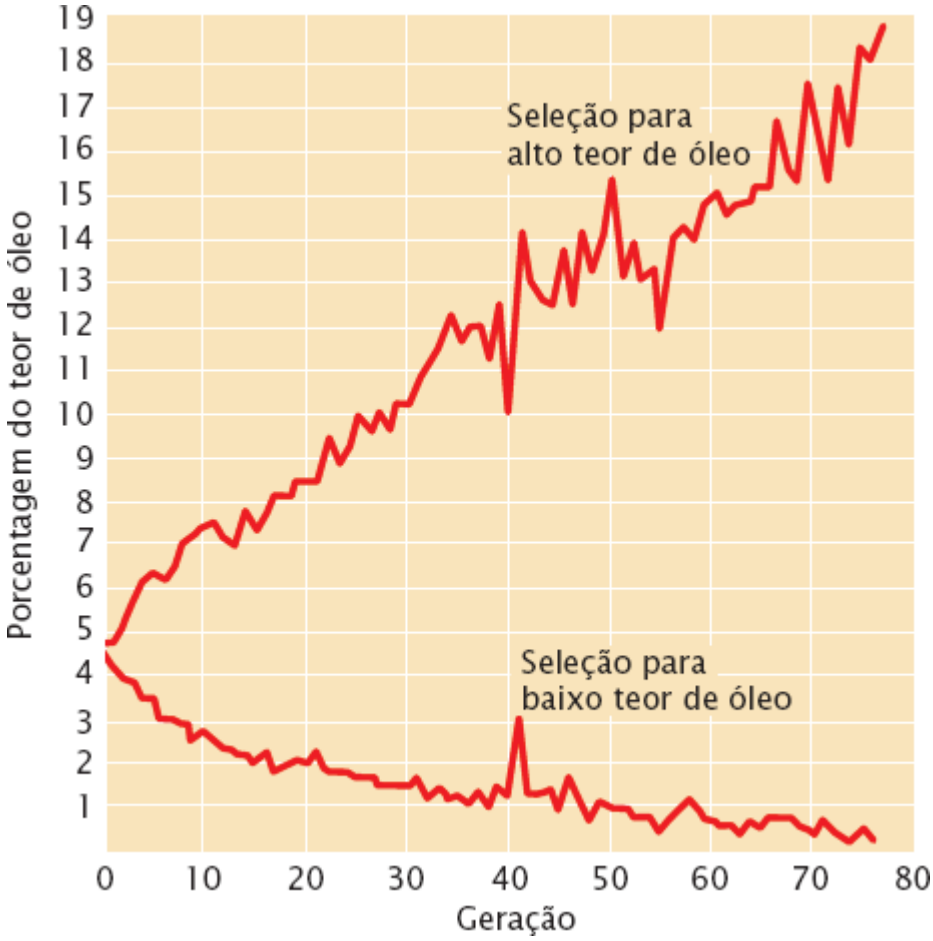


Figura 24.22 Em um experimento a longo prazo de resposta à seleção, a seleção para o teor de óleo no milho aumentou o teor do óleo em uma linhagem para cerca de 20%, enquanto quase eliminou em outra linhagem.

Limites para a resposta à seleção

Quando uma característica é selecionada por muitas gerações, a resposta acaba estabilizando e a característica não responde mais à seleção (**Figura 24.23**). Uma possível resposta para esta estabilização é que o esgotamento da variação genética na população; em algum ponto, todos os indivíduos na população se tornaram homozigotos para os alelos que codificam o traço selecionado. Quando não existe mais variação genética aditiva, a herdabilidade é igual a zero e não pode mais ocorrer resposta à seleção.

Em alguns casos, a resposta à seleção pode estabilizar mesmo enquanto ainda permanece variação genética na população. Esta estabilização ocorre porque a seleção natural se contrapõe à mudança na característica. A resposta à seleção para o corpo pequeno nos camundongos, por exemplo, consequentemente estabiliza porque os menores animais são estéreis e não conseguem transmitir seus genes de pequeno tamanho de corpo. Neste caso, a seleção artificial para o tamanho pequeno é contraposta pela seleção natural para fertilidade e a população não consegue mais responder à seleção artificial.

Respostas correlacionadas

Muitas vezes, quando um traço específico é selecionado, outros traços mudam ao mesmo tempo. Este tipo de resposta associada se deve ao fato de que os traços são codificados pelos mesmos genes.

Duas ou mais características estão, com frequência, correlacionadas. Por exemplo, em muitas plantas, o tamanho da planta e o número de sementes produzidas exibem uma correlação positiva: plantas maiores, em média, produzem mais sementes do que as plantas menores. Esta correlação é uma **correlação fenotípica** porque a associação é entre dois fenótipos no mesmo indivíduo. As correlações fenotípicas podem ocorrer por correlações ambientais ou genéticas. As correlações ambientais se referem a duas ou mais características influenciadas pelo mesmo fator ambiental. A disponibilidade de umidade, por exemplo, pode afetar o tamanho da planta e o número de sementes produzidas pela planta. As plantas que crescem em ambientes com muita água são maiores e produzem muitas sementes, enquanto as plantas que crescem em ambientes com água limitada são menores e têm menos sementes.

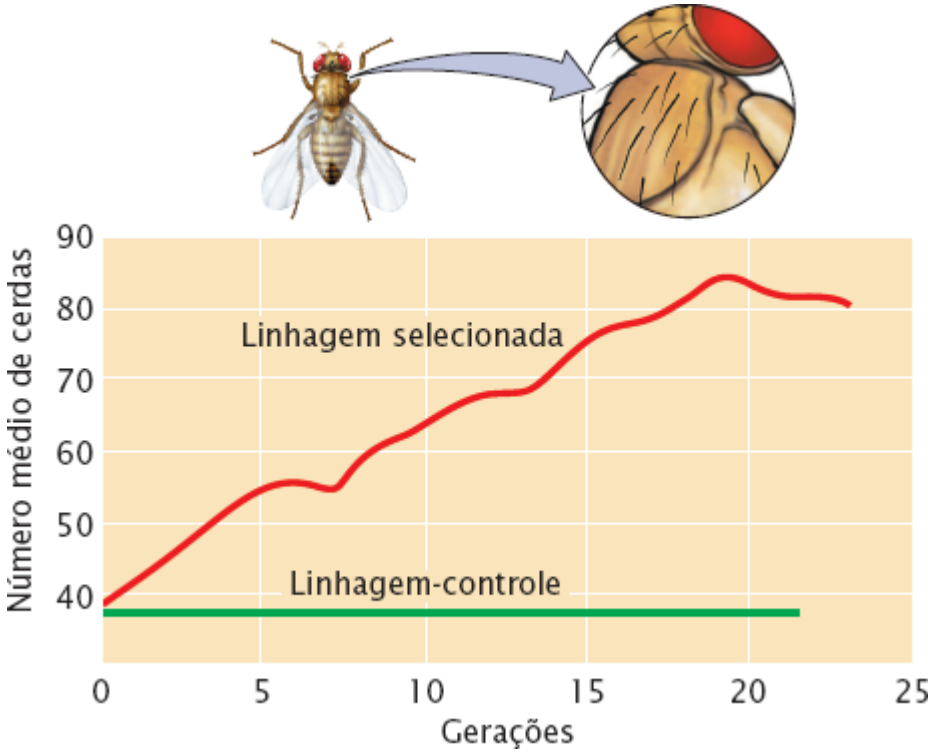


Figura 24.23 A resposta de uma população à seleção com frequência estabiliza em algum momento. Em um experimento de resposta à seleção para número maior de cerdas de pelos abdominais nas moscas-da-fruta fêmeas, o número de cerdas aumentou de forma constante por cerca de 20 gerações e então estabilizou-se.

Por outro lado, uma correlação fenotípica pode surgir a partir de uma **correlação genética**, o que significa que os genes que afetam duas características estão associados. A causa genética primária das correlações fenotípicas é a pleiotropia, que ocorre pelo efeito de um gene em duas ou mais características (ver *Capítulo 5*). Nos humanos, por exemplo, muitas estruturas corporais respondem ao hormônio do crescimento (GH), e existem genes que afetam a quantidade de hormônio de crescimento secretado pela hipófise. Pessoas com alguns genes produzem níveis elevados de GH, o que aumenta a altura e o tamanho da mão. Outras têm genes que produzem níveis menores do hormônio do crescimento, o que leva a menor altura e mãos menores. A altura e o tamanho das mãos estão, portanto, fenotipicamente correlacionadas nos humanos e esta correlação se deve a uma

correlação genética, o fato de que ambas as características são afetadas pelos genes que controlam a quantidade do GH. Do ponto de vista genético, a altura e o tamanho das mãos são a mesma característica porque eles são a manifestação fenotípica de um único conjunto de genes. Quando duas características são influenciadas pelos mesmos genes, elas estão geneticamente correlacionadas.

As correlações genéticas são muito comuns (**Quadro 24.3**) e podem ser positivas ou negativas. Uma correlação genética positiva entre duas características significa que os genes que provocam um aumento em uma característica também produzem um aumento na outra característica. O comprimento do tórax e da asa na *Drosophila* estão positivamente correlacionados porque os genes que aumentam o comprimento do tórax também aumentam o comprimento da asa. Uma correlação genética negativa entre duas características significa que os genes que provocam um aumento em uma característica também produzem uma redução na outra característica. A produção de leite e a porcentagem de gordura estão negativamente correlacionadas no gado: genes que favorecem maior produção de leite resultam em leite com menor porcentagem de gordura (nata).

Quadro 24.3 Correlações genéticas em vários organismos.

Organismo	Características	Correlação genética
Gado	Produção de leite e porcentagem de gordura	–0,38
Porco	Ganho de peso e espessura de gordura na parte traseira do corpo	0,13
	Ganho de peso e eficiência	0,69
Galinha	Peso do corpo e peso do ovo	0,42
	Peso do corpo e produção de ovos	–0,17
	Peso do ovo e produção de ovos	–0,31
Camundongo	Peso do corpo e comprimento da cauda	0,29
Mosca-da-fruta	Número de cerdas abdominais e número de cerdas esternopleurais	0,41

Fonte: De D. S. Falconer and T. F. C. Mackay. Introduction to Quantitative Genetics (Pearson, New York, 1996), p. 314.

As correlações genéticas são importantes na reprodução de animais e de plantas porque produzem uma resposta correlata à seleção, o que significa que, quando uma característica é adicionada, as características geneticamente correlatas também mudam. As respostas correlatas à seleção se devem ao fato de que ambas as características são influenciadas pelos mesmos genes; a seleção para uma característica provoca uma mudança nos genes que afetam aquela característica e estes genes também afetam a segunda característica, fazendo-a mudar ao mesmo tempo. As respostas correlacionadas podem ser bem indesejáveis e podem limitar a capacidade de alterar uma característica pela seleção. De 1944 a 1964, os perus domesticados foram submetidos a intensa seleção para taxa de crescimento e tamanho do corpo. Ao mesmo tempo, a fertilidade, a produção de ovos e a eclodibilidade dos ovos declinaram. Estas respostas correlatas se devem a correlações genéticas negativas entre tamanho do corpo e fertilidade; estas correlações genéticas acabaram limitando a taxa de crescimento dos perus à seleção. As correlações genéticas também limitam a capacidade de populações naturais em responder à seleção no ambiente selvagem e se adaptar ao seu ambiente. Resolva o Problema 45

Conceitos

As correlações genéticas surgem a partir da pleiotropia. Quando duas características estão geneticamente correlacionadas, a seleção para uma característica produzirá uma resposta correlata na outra característica.

Checagem dos conceitos 8

Em um rebanho de gado leiteiro, a produção de leite e a porcentagem de gordura (nata) exibem uma correlação genética de –0,38. Se a maior produção de leite for selecionada neste rebanho, qual será o efeito na porcentagem de gordura (nata)?

Resumo dos conceitos

- A genética quantitativa se dedica à herança de características complexas cujos fenótipos variam de forma contínua. Para muitas características quantitativas, a relação entre genótipo e fenótipo é complexa porque muitos genes e fatores ambientais influenciam uma característica
- Os genes individuais que influenciam uma característica poligênica seguem os mesmos princípios mendelianos que governam as características descontínuas, mas como muitos genes participam, as razões esperadas de fenótipos são pouco claras
- Uma população é o grupo de interesse e uma amostra é um subconjunto da população usado para descrevê-la
- Uma distribuição de frequência, na qual os fenótipos são representados em um eixo e o número de indivíduos com cada fenótipo é representado no outro eixo, é um meio conveniente para resumir os fenótipos encontrados em um grupo de indivíduos
- A média e a variância fornecem informações importantes sobre uma distribuição: a média fornece a localização central da distribuição e a variância informa como o fenótipo varia em um grupo
- O coeficiente de correlação mede a direção e a força da associação entre duas variáveis. A regressão pode ser usada para prever o valor de uma variável com base no valor de uma variável correlacionada
- A variância fenotípica em uma característica pode ser dividida em componentes que são variância genética aditiva, variância genética por dominância, variância da interação genética, variância ambiental e variância da interação genética ambiental
- A herdabilidade de sentido amplo é a proporção de variância fenotípica devido à variância genética, a herdabilidade de sentido restrito é a proporção de variância fenotípica devido à variância genética aditiva
- A herdabilidade informa apenas o grau no qual a variação em uma característica resulta em diferenças genéticas. A herdabilidade baseia-se nas variâncias existentes em um grupo de indivíduos e um indivíduo não tem herdabilidade. A herdabilidade de uma característica varia entre as populações e os ambientes. Mesmo que a herdabilidade para uma característica seja alta, a característica ainda pode ser alterada por mudanças no ambiente. A herdabilidade não fornece informações sobre a natureza das diferenças da população em uma característica
- Os *loci* de traço quantitativo (QTL) são segmentos cromossômicos com genes que controlam características poligênicas. Os QTLs podem ser mapeados ao examinar a associação entre a herança de uma característica quantitativa e a herança de marcadores genéticos. Os genes que influenciam traços quantitativos também podem ser localizados com o uso de estudos de associação genômica ampla
- O quanto uma característica quantitativa é modificada em uma única geração quando sujeita à seleção (a resposta à seleção) está diretamente relacionado à seleção diferencial e à herdabilidade de sentido restrito
- Pode existir correlação genética quando o mesmo gene afeta duas ou mais características (pleiotropia). As correlações genéticas produzem respostas correlacionadas à seleção.

Termos importantes

Amostra
Característica com limiar
Característica merística
Coeficiente de correlação
Coeficiente de regressão
Correlação
Correlação fenotípica
Correlação genética
Desvio padrão
Distribuição de frequência
Distribuição normal
Genética quantitativa
Herdabilidade
Herdabilidade de sentido amplo
Herdabilidade de sentido restrito
Herdabilidade realizada
Loci de traço quantitativo (QTL)
Média
População
Regressão

Resposta à seleção
Seleção artificial
Seleção diferencial
Seleção natural
Variância
Variância ambiental
Variância da interação gênica
Variância da interação genética ambiental
Variância fenotípica
Variância genética
Variância genética aditiva
Variância genética por dominância

Respostas da Checagem dos conceitos

1. Cruzar dois indivíduos que são, cada um, homozigotos para diferentes genes que afetam os traços e, depois, intercruzar os descendentes F_1 resultantes para produzir F_2. Determinar qual proporção de descendentes F_2 lembra um dos homozigotos originais na geração P. Esta proporção deve ser $(1/4)^n$, na qual n é igual ao número de *loci* com um par segregante de alelos que influenciam a característica.
2. d.
3. d.
4. b.
5. a.
6. Ele indica que cerca de 40% das diferenças na pressão arterial de afro-americanos em Detroit se devem a diferenças genéticas aditivas. Não fornece informações sobre a herdabilidade da pressão arterial em outros grupos de pessoas nem indica nada sobre a natureza das diferenças pressóricas em afro-americanos em Detroit e pessoas de outros grupos.
7. 0,2.
8. A porcentagem de gordura da manteiga reduzirá.

Problemas desenvolvidos

Problema 1

O peso da semente em uma espécie particular de planta é determinado por pares de alelos em dois *loci* (a^+a^- e b^+b^-) que são aditivos e com efeitos iguais. As plantas com genótipo $a^-a^-\ b^-b^-$ têm sementes com peso médio de 1 g, enquanto as plantas com genótipo $a^+a^+\ b^+b^+$ têm sementes com peso médio de 3,4 g. Uma planta com genótipo $a^-a^-\ b^-b^-$ é cruzada com uma planta com genótipo $a^+a^+\ b^+b^+$.

a. Qual é o peso previsto das sementes dos descendentes deste cruzamento?

b. Se as plantas forem intercruzadas, quais são os pesos de sementes esperados e as proporções das plantas F_2?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

- O peso previsto de sementes dos descendentes
- Os pesos esperados das sementes e suas proporções nas plantas F_2.

Quais informações são fornecidas para solucionar o problema?

- O peso da semente é determinado por pares de alelos em dois *loci* (a^+a^- e b^+b^-)
- Os alelos são aditivos e com efeitos iguais
- As plantas com genótipo $a^-a^-\ b^-b^-$ têm sementes com peso médio de 1 g
- As plantas com genótipo $a^+a^+\ b^+b^+$ têm sementes com peso médio de 3,4 g
- Uma planta com genótipo $a^-a^-\ b^-b^-$ é cruzada com uma planta com genótipo $a^+a^+\ b^+b^+$.

Para a solução deste problema, revise:
Herança poligênica, na Seção 24.1.

Etapas para a solução

A diferença no peso médio da semente dos dois genótipos parentais é 3,4 g – 1 g = 2,4 g. Estes dois genótipos diferem em quatro genes, então cada diferença de gene contribui com um adicional de 2,4 g/4 = 0,6 g no peso para 1 g da planta ($a^-a^-\ b^-b^-$) que não tem nenhum dos genes que contribuem.

Dica: Como os alelos são iguais e aditivos, cada alelo contribui com a mesma quantidade para a altura.

O cruzamento entre os dois genótipos homozigotos produz os descendentes F_1 e F_2 apresentados na página a seguir.

a. Os F_1 são heterozigotos em ambos *loci* ($a^+a^-\ b^+b^-$) e têm dois genes que contribuem com 0,6 g cada para o peso de 1 g de uma planta sem genes contribuintes. Portanto, as sementes devem ter em média 1 g + 2(0,6 g) = 2,2 g.

b. F_2 terá os seguintes fenótipos e proporções: 1/16 1 g; 4/16 1,6 g; 6/16 2,2 g; 4/16 2,8 g; e 1/16 3,4 g.

P $a^-a^-\ b^-b^-$ (1 g) × $a^+a^+\ b^+b^+$ (3,4 g)

F_1 $a^+a^-\ b^+b^-$ (2,2 g)

Lembrete: A probabilidade de cada genótipo de dois *loci* pode ser determinada ao multiplicar a probabilidade dos genótipos de *locus* único.

	Genótipo	Probabilidade		Número de genes que contribuem	Peso médio da semente
	$a^-a^-\ b^-b^-$	$1/4 \times 1/4 = 1/16$		0	$1\text{ g} + (0 \times 0{,}6\text{ g}) = 1\text{ g}$
	$a^+a^-\ b^-b^-$	$1/2 \times 1/4 = 1/8$	$2/8 = 4/16$	1	$1\text{ g} + (1 \times 0{,}6\text{ g}) = 1.6\text{ g}$
	$a^-a^-\ b^+b^-$	$1/4 \times 1/2 = 1/8$			
F_2	$a^+a^+\ b^-b^-$	$1/4 \times 1/4 = 1/16$			
	$a^-a^-\ b^+b^+$	$1/4 \times 1/4 = 1/16$	$2/16 + 1/4 = 6/16$	2	$1\text{ g} + (2 \times 0{,}6\text{ g}) = 2.2\text{ g}$
	$a^+a^-\ b^+b^-$	$1/2 \times 1/2 = 1/4$			
	$a^+a^+\ b^+b^-$	$1/4 \times 1/2 = 1/8$	$2/8 = 4/16$	3	$1\text{ g} + (3 \times 0{,}6\text{ g}) = 2.8\text{ g}$
	$a^+a^-\ b^+b^+$	$1/4 \times 1/2 = 1/8$			
	$a^+a^+\ b^+b^+$	$1/4 \times 1/4 = 1/16$		4	$1\text{ g} + (4 \times 0{,}6\text{ g}) = 3.4\text{ g}$

Problema 2

Um fazendeiro cria coelhos. O peso corporal médio na sua população de coelhos é 3 kg. Ele seleciona os 10 maiores coelhos na sua população, cujo peso corporal médio é 4 kg, e faz o cruzamento entre eles. Se a herdabilidade do peso corporal na população de coelhos é 0,7, qual é o peso corporal esperado dos descendentes dos coelhos selecionados?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
O peso esperado dos descendentes dos coelhos selecionados.

Quais informações são fornecidas para solucionar o problema?

- O peso corporal médio na população é 3 kg
- O peso corporal médio dos coelhos selecionados é 4 kg

- A herdabilidade do peso corporal é 0,7.

Para a solução deste problema, revise:
Como prever a resposta à seleção, na Seção 24.4.

Etapas para a solução

O fazendeiro fez um experimento de resposta à seleção. A seleção diferencial é igual à diferença nos pesos médios dos coelhos selecionados e na população inteira: 4 kg – 3 kg = 1 kg. A herdabilidade de sentido restrito é 0,7; então a resposta esperada à seleção é: $R = h^2 \times S$ = 0,7 × 1 kg = 0,7 kg. Este valor é o aumento do peso que é esperado nos descendentes dos genitores selecionados, então o peso médio dos descendentes é: 3 kg + 0,7 kg = 3,7 kg.

Lembrete: A resposta à seleção é igual à seleção diferencial multiplicada pela herdabilidade de sentido restrito.

Questões de compreensão

Seção 24.1

1. Qual a diferença entre uma característica quantitativa e uma característica descontínua?
2. Explique resumidamente por que a relação entre genótipo e fenótipo é frequentemente complexa para as características quantitativas.
3. Por que as características poligênicas têm muitos fenótipos?

Seção 24.2

4. Explique a relação entre uma população e uma amostra. Quais características uma amostra deve ter para ser representativa da população?
5. Quais informações a média e a variância fornecem sobre a distribuição?
6. Qual a relação entre desvio padrão e variância?
7. Que informação o coeficiente de correlação fornece sobre a associação entre duas variáveis.
8. O que é regressão? Como é usada?

Seção 24.3

9. Liste todos os componentes que contribuem para a variância fenotípica e defina cada componente.
10. Qual a diferença entre herdabilidade de sentido amplo e de sentido restrito?
11. Descreva resumidamente algumas das formas nas quais a herdabilidade pode ser calculada.
12. Descreva resumidamente os mal-entendidos e usos errados do conceito de herdabilidade.
13. Explique resumidamente como os genes que afetam uma característica poligênica são localizados com o uso de mapeamento por QTL.

Seção 24.4

14. Qual é a relação da resposta à seleção à herdabilidade de sentido restrito e à seleção diferencial? Que informação a resposta à seleção fornece?
15. Por que a resposta à seleção frequentemente estabiliza após muitas gerações de seleção?

Questões e problemas aplicados

Seção 24.1

*16. Para cada uma das características a seguir, indique se ela seria considerada uma característica descontínua ou uma característica quantitativa. Justifique resumidamente sua resposta.
    a. A cor do grão em uma cepa de trigo, no qual dois alelos codominantes se separam em um *locus* único determinam a cor. Assim, existem três fenótipos presentes nesta cepa: branco, vermelho-claro e vermelho médio.
    b. O peso corporal em uma família de Labradores retrievers. Um alelo autossômico recessivo que causa nanismo é

encontrado nesta família. Dois fenótipos são reconhecidos: nanismo (menos de 13 kg) e normal (mais de 23 kg).

c. Ocorrência ou não de hanseníase. A suscetibilidade à hanseníase é determinada por múltiplos genes e numerosos fatores ambientais.

d. Número de dedos da pata dos porquinhos-da-índia, que é influenciado por muitos *loci*.

e. Número de dedos nas mãos dos humanos. Dedos extras (mais de cinco) são causados por um alelo autossômico dominante.

*17. Considere que o peso da planta é determinado por um par de alelos em cada *loci* de segregação independente (*A* e *a*, *B* e *b*) que são aditivos em seus efeitos. Agora considere que cada alelo representado por uma letra maiúscula contribui 4 g para o peso e que cada alelo representado pela letra minúscula contribui 1 g para o peso.

a. Se uma planta com genótipo *AA BB* é cruzada com uma planta com genótipo *aa bb*, que pesos são esperados nos descendentes F_1?

b. Qual a distribuição de peso esperada nos descendentes F_2?

18. Considere que três *loci*, cada um com dois alelos (*A* e *a*, *B* e *b*, *C* e *c*), determinam as diferenças na altura entre duas cepas homozigotas de uma planta. Estes genes são aditivos e iguais nos seus efeitos sobre a altura da planta. Uma cepa (*aa bb cc*) tem 10 cm de altura. A outra cepa (*AA BB CC*) tem 22 cm de altura. As duas cepas são cruzadas e os F_1 resultantes são intercruzados para produzir os descendentes F_2. Indique os fenótipos e proporções esperadas dos descendentes F_2.

*19. Um fazendeiro tem duas variedades homozigotas de tomates. Uma variedade, chamada Little Pete, tem frutos com média de apenas 2 cm de diâmetro. A outra variedade, Big Boy, tem frutos com uma média de fantásticos 14 cm de diâmetro. O fazendeiro cruza Little Pete e Big Boy, então intercruza F_1 para produzir os descendentes F_2. Ele cultiva 200 tomates F_2 e não encontra nenhum descendente F_2 que produza frutos tão pequenos como Little Pete ou tão grandes como Big Boy. Se assumirmos que as diferenças no tamanho do fruto destas variedades são produzidas por genes com efeitos iguais e aditivos, o que podemos concluir sobre o número mínimo de *loci* com pares de alelos que determinam as diferenças no tamanho do fruto das duas variedades?

20. O tamanho da semente em uma planta é uma característica poligênica. Uma cultivadora cruza duas variedades de linhagem pura da planta e mede o tamanho da semente nos descendentes F_1. Então faz o retrocruzamento das plantas F_1 de uma das variedades parentais e mede o tamanho da semente nos descendentes deste cruzamento. Ela descobre que o tamanho da semente nos descendentes do retrocruzamento tem uma variância maior que o tamanho da semente nos descendentes F_1. Explique por que os descendentes do retrocruzamento são mais variáveis.

Seção 24.2

21. Os dados a seguir são os números de dedos por pata em 25 porquinhos-da-índia. Construa uma distribuição de frequência para estes dados.

4, 4, 4, 5, 3, 4, 3, 4, 4, 5, 4, 4, 3, 2, 4, 4, 5, 6, 4, 4, 3, 4, 4, 4, 5

*22. Dez estudantes do sexo masculino de Harvard foram pesados em 1916. Seus pesos são fornecidos em quilogramas. Calcule a média, a variância e o desvio padrão desses pesos.

51, 69, 69, 57, 61, 57, 75, 105, 69, 63

23. Em uma população de girinos, o coeficiente de correlação para o tamanho na metamorfose e tempo necessário para a metamorfose é –0,74. Com base nesta correlação, que conclusões você pode tirar sobre os tamanhos relativos dos girinos que sofrem metamorfose rapidamente e os que sofrem metamorfose mais lentamente?

24. Um pesquisador, ao estudar o consumo de bebidas alcoólicas nas cidades da América do Norte, encontra uma correlação positiva e significativa entre o número de pastores batistas e o consumo de bebidas alcoólicas. É válido o pesquisador concluir que os batistas estão consumindo mais etanol? Por que sim ou por que não?

*25. Foram medidos o peso corporal e comprimento de seis peixes-mosquito; essas medidas estão descritas a seguir. Calcule o coeficiente de correlação para o peso e o comprimento nestes peixes.

Peso úmido (g)	Comprimento (m)
115	18
130	19
210	22
110	17

140	20
185	21

26. As alturas das mães e filhas estão descritas a seguir:

Altura da mãe (cm)	Altura da filha (cm)
162,5	167,6
165,1	167,6
167,6	172,7
162,5	165,1
160,0	165,1
160,0	157,4
150,0	157,4
157,4	162,5
155,0	160,0
152,4	157,4

a. Calcule o coeficiente de correlação para as alturas das mães e filhas.
b. Usando a regressão, preveja a altura esperada de uma filha cuja mãe tenha 170,1 cm de altura.

Seção 24.3

*27. A variação fenotípica no comprimento da cauda dos camundongos tem os seguintes componentes:

Variância genética aditiva (V_A) = 0,5
Variância genética por dominância (V_D) = 0,3
Variância da interação genética (V_I) = 0,1
Variância ambiental (V_E) = 0,4
Variância da interação genética ambiental (V_{GE}) = 0,0

a. Qual é a herdabilidade de sentido restrito do comprimento da cauda?
b. Qual é a herdabilidade de sentido amplo do comprimento da cauda?

28. A herdabilidade de sentido restrito do comprimento da orelha nos coelhos Reno é 0,4. A variância fenotípica (V_P) é 0,8 e a variância ambiental (V_E) é 0,2. Qual é a variância genética aditiva (V_A) para o comprimento das orelhas destes coelhos?

29. Considere que o comprimento da orelha humana é influenciado por múltiplos fatores genéticos e ambientais. Suponha que você mediu o comprimento da orelha de três grupos de pessoas, nas quais o grupo A tinha cinco pessoas não próximas, grupo B tinha cinco irmãos e grupo C tinha cinco primos em primeiro grau.
a. Com a hipótese de que o ambiente para cada grupo é semelhante, qual grupo deve ter a maior variância fenotípica? Explique por quê.
b. É válido considerar que a variância ambiental para cada grupo é semelhante? Explique sua resposta.

30. Uma característica tem uma herdabilidade de sentido restrito de 0,6.
a. Se a variância da dominância (V_D) aumentar e todos os outros componentes de variância permanecerem iguais, o que acontecerá com a herdabilidade de sentido restrito? Ela vai aumentar, diminuir, ou permanecer a mesma? Explique.
b. O que acontecerá com a herdabilidade de sentido amplo? Explique.
c. Se a variância ambiental (V_E) aumenta e todos os outros componentes de variância permanecem os mesmos, o que acontecerá com a herdabilidade de sentido restrito? Explique.
d. O que acontecerá com a herdabilidade de sentido amplo? Explique.

31. A cor da flor nas variedades de ervilhas estudadas por Mendel é controlada por alelos em um único *locus*. Um grupo de ervilhas homozigotas para flores de coloração roxa é cultivado. O estudo cuidadoso das plantas revela que todas suas flores são roxas, mas existe alguma variabilidade na intensidade da cor púrpura. Qual seria a herdabilidade estimada para esta variação na cor da flor? Explique sua resposta.

*32. Uma estudante de graduação está estudando uma população de tremoços-azuis (*Lupinus*) ao longo de uma estrada. As plantas nesta população são geneticamente variáveis. Ela conta as sementes produzidas por 100 plantas e mensura a média e variância do número de sementes. A variância é 20. Ao selecionar uma planta, a estudante tira ramos e os cultiva na estufa, produzindo muitos clones geneticamente idênticos da mesma planta. Ela então transplanta estes clones para a

população da estrada, os deixa crescerem por 1 ano e então conta o número de sementes produzidas por cada uma das plantas clones. A estudante descobre que a variância no número de sementes dessas plantas clonadas é 5. A partir da variância fenotípica das plantas geneticamente variáveis e geneticamente idênticas, ela calcula a herdabilidade de sentido amplo.

a. Qual é a herdabilidade de sentido amplo do número de sementes para a população de tremoço-azul na estrada?
b. O que poderia fazer com que esta estimativa de herdabilidade não fosse exata?

33. Muitos pesquisadores estimaram a herdabilidade de traços humanos por meio de comparação dos coeficientes de correlação de gêmeos monozigóticos e dizigóticos (ver *Capítulo 6*). Uma das ideias ao usar este método é que gêmeos monozigóticos experimentam ambientes que não são mais semelhantes do que os outros experimentados pelos gêmeos dizigóticos. Como esta ideia poderia ser violada? Indique alguns exemplos específicos de como os ambientes de gêmeos monozigóticos poderiam ser mais semelhantes do que os ambientes de gêmeos dizigóticos.

34. Que conclusão você pode tirar da Figura 24.18 sobre a proporção de variação fenotípica na largura da concha que se deve a diferenças genéticas? Explique sua resposta.

*35. Um pesquisador em genética determina que a herdabilidade de sentido amplo da altura nos graduandos da Southwestern University é 0,9. Qual das conclusões a seguir seria válida? Explique sua resposta.

a. Como Sally é uma estudante da Southwestern University, 10% da sua altura são determinados por fatores não genéticos.
b. Noventa por cento da variação da altura em todos os universitários nos EUA se devem a diferenças genéticas.
c. Noventa por cento da altura dos estudantes da Southwestern University são determinados pelos genes.
d. Dez por cento da variação na altura nos estudantes da Southwestern University são determinados pela variação em fatores não genéticos.
e. Como a herdabilidade da altura desses estudantes é tão elevada, qualquer mudança no ambiente dos estudantes terá efeito mínimo na sua altura.

*36. O comprimento da articulação média do dedo indicador direito foi medido em 10 conjuntos de pais e seus descendentes adultos. Os comprimentos médios parentais e os comprimentos médios dos descendentes para cada família estão listados no quadro a seguir. Calcule o coeficiente de regressão para regressão do comprimento médio dos descendentes contra o comprimento médio parental e estime a herdabilidade de sentido restrito dessa característica.

Comprimento médio parental (mm)	Comprimento médio dos descendentes (mm)
30	31
35	36
28	31
33	35
26	27
32	30
31	34
29	28
40	38
33	34

37. Considere que na Figura 24.14, *x* é igual ao fenótipo médio dos genitores e *y* é igual ao fenótipo médio dos descendentes. Qual linha representa a maior herdabilidade? Explique sua resposta.

ANÁLISE DE DADOS

38. A *Drosophila buzzatii* é uma mosca-da-fruta que se alimenta de frutas em putrefação de cactos na Austrália. Timothy Prout e Stuart Barker calcularam a herdabilidade do tamanho do corpo, medido pelo comprimento do tórax, de uma população natural de *D. buzzatii* criada na natureza e de uma população de *D. buzzatii* coletada na natureza, mas criada no laboratório (T. Prout and J. S.F. Barker. 1989. *Genetics* 123:803-813). Eles encontraram as seguintes herdabilidades:

População	Herdabilidade do tamanho do corpo (± erro padrão)
População selvagem	0,0595 ± 0,0123
População criada em laboratório	0,3770 ± 0,0203

Você acha que a herdabilidade calculada na população criada em laboratório é maior que a medida na população criada na natureza?

39. Sr. Jones é um criador de porcos. Por muitos anos, ele alimentou seus porcos com os restos de alimentos da cafeteria da universidade local, que é conhecida por ter pouca proteína, deficiente em vitaminas e, obviamente, sem gosto. Entretanto, a comida é de graça, e seus porcos não reclamam. Um dia, um vendedor de uma indústria alimentícia visitou Sr. Jones. O vendedor afirmou que sua empresa vende uma ração nova, com alto teor de proteína e enriquecida com vitaminas que melhora o ganho de peso nos porcos. Embora a ração seja cara, o vendedor afirma que o ganho maior de peso dos porcos pagará pelo custo da ração, aumentando o lucro de Sr. Jones. Sr. Jones responde que frequentou uma turma de genética na universidade e fez alguns experimentos genéticos nos seus porcos, especificamente, e calculou a herdabilidade de sentido restrito do ganho de peso para seus porcos, que era 0,98. Ele disse que este valor de herdabilidade indica que 98% da variância no ganho de peso entre seus porcos são determinados por diferenças genéticas e, portanto, a nova ração para porcos terá pouco efeito no crescimento dos seus animais. E concluiu que o alimento seria um desperdício de dinheiro. O vendedor não argumentou a estimativa de herdabilidade do Sr. Jones, mas ainda afirma que a nova ração pode melhorar de forma significativa o ganho de peso nos seus porcos. Quem está certo e por quê?

Seção 24.4

*40. Joe está criando baratas no seu dormitório. Ele descobriu que o comprimento médio das asas de sua população de baratas é 4 cm. Ele escolheu seis baratas com as maiores asas; o comprimento médio entre estas baratas selecionadas é 10 cm. Joe cruzou estas baratas selecionadas. A partir de estudos anteriores, ele sabe que a herdabilidade de sentido restrito para o comprimento da asa na sua população é 0,6.
   a. Calcule a seleção diferencial e a resposta esperada à seleção para o comprimento das asas nestas baratas.
   b. Qual deve ser o comprimento médio da asa nos descendentes das baratas selecionadas?

41. Três características do gado para corte, peso corporal, teor de gordura e maciez da carne, são determinadas e os seguintes componentes de variância são estimados:

	Peso corporal	Teor de gordura	Maciez da carne
V_A	22	45	12
V_D	10	25	5
V_I	3	8	2
V_E	42	64	8
V_{GE}	0	0	1

Nesta população, qual característica responderia melhor à seleção? Explique sua resposta.

42. Uma fazendeira determina que a quantidade média de lã produzida por um carneiro no seu rebanho é de 22 kg por ano. Em uma tentativa de aumentar a produção de lã de seu rebanho, ela seleciona cinco machos e cinco fêmeas com a maior produção de lã; a quantidade média de lá produzida por carneiro destes selecionados é de 30 kg. Ela promove o intercruzamento desses carneiros e descobre que a produção média de lã dos descendentes deles é de 28 kg. Qual é a herdabilidade de sentido restrito para a produção de lã nos carneiros no rebanho da fazendeira?

43. Um criador de morangos determina que o peso médio de cada morango produzido por suas plantas no seu jardim é 2 g. Ele seleciona as 10 plantas que produzem os maiores morangos; o peso médio dos morangos dessas plantas selecionadas é de 6 g. A seguir, ele promove o intercruzamento dessas plantas. Os descendentes delas produzem morangos com 5 g. Se o criador tivesse selecionado as plantas que produzem um morango de peso médio de 4 g, qual seria o peso previsto dos morangos produzidos pelos descendentes das mesmas?

44. A resposta à seleção diminuiu na cepa de milho selecionada para o alto teor de óleo apresentado na Figura 24.22? O que essa observação sugere sobre a variação genética na cepa selecionada quanto ao alto teor de óleo?

*45. A herdabilidade de sentido restrito do comprimento de asa em uma população de *Drosophila melanogaster* é 0,8. A herdabilidade de sentido restrito da largura da cabeça na mesma população é 0,9. A correlação genética entre comprimento da asa e largura da cabeça é –0,86. Se um geneticista selecionar o comprimento maior de asas nestas moscas, o que acontecerá com a largura da cabeça?

46. Os porcos foram domesticados a partir dos porcos-do-mato. Você esperaria encontrar maior herdabilidade para peso nos porcos domesticados ou nos porcos-do-mato? Explique sua resposta.

Questões desafiadoras

Seção 24.1

47. O transtorno bipolar é uma doença psiquiátrica com forte base hereditária, mas o modo exato de herança é desconhecido. A

pesquisa mostrou que irmãos de pacientes com transtorno bipolar têm maior probabilidade de desenvolver a doença do que irmãos de pessoas não afetadas. Os achados de um estudo demonstraram que a razão irmãos bipolares:irmãs bipolares é maior quando o paciente é um homem do que quando a paciente é mulher. Em outras palavras, relativamente mais irmãos de pacientes bipolares também apresentam a doença quando o paciente é homem do que quando a paciente é mulher. O que essa observação sugere sobre a herança do transtorno bipolar?

Seção 24.3

48. Exploramos algumas das dificuldades em separar os componentes genéticos e ambientais das características comportamentais dos humanos. Considerando estas dificuldades e o que você sabe sobre cálculo de herdabilidade, proponha um modelo experimental para medir com exatidão a herdabilidade da capacidade musical.
49. Um estudante que acabou de aprender sobre genética quantitativa diz: “As estimativas da herdabilidade são inúteis! Elas não dizem nada sobre os genes que afetam uma característica. Não fornecem informações sobre os tipos de descendentes esperados de um cruzamento. As estimativas de herdabilidade medidas em uma população não podem ser usadas para outras populações, então não fornecem nenhuma informação geral sobre o quanto uma característica é determinada pela genética. A herdabilidade não é útil, apenas faz os estudantes sofrerem durante os exames”. Como você responderia a esta afirmação? O estudante está correto? Qual é o valor desse coeficiente genético e por que os geneticistas se importam em calculá-lo?

Seção 24.4

50. Eugene Eisen selecionou uma ninhada de 12 dias de peso maior (peso total de uma ninhada de 12 dias após o nascimento) em uma população de camundongos (E. J. Eisen. 1972. *Genetics* 72:129-142). O peso da ninhada de 12 dias da população aumentou de forma constante, mas então estabilizou-se após cerca de 17 gerações. Na geração 17, Eisen pegou uma família de camundongos da população selecionada e inverteu o procedimento de seleção: neste grupo, ele selecionou o peso menor da ninhada de 12 dias. Este grupo respondeu imediatamente à seleção reduzida; o peso da ninhada de 12 dias caiu 4,8 g em uma geração e caiu 7,3 g após cinco gerações. Com base nos resultados da seleção reversa, qual é a explicação mais plausível para a estabilização do peso da ninhada de 12 dias na população original?

25 Genética de Populações

Resgate genético do carneiro-das-montanhas

O carneiro-das-montanhas (*Ovis canadensis*) é um dos animais mais espetaculares da América do Norte, caracterizado pelos magníficos chifres do macho que se curvam para trás sobre as orelhas, enrolando para baixo e para trás perto da face. Há duzentos anos, eram numerosos em toda região oeste da América do Norte, indo do México até sul de Alberta e do Colorado para Califórnia. Meriwether Lewis e William Clark relataram numerosos avistamentos destes lindos animais na sua expedição pelo Oeste dos EUA de 1804 a 1806. Antes de 1900, existiam em torno de 2 milhões de carneiros-das-montanhas na América do Norte.

Infelizmente, a colonização do Oeste pelos europeus não foi boa para os carneiros. Começando no final dos anos 1800, a caça, a perda de *habitat*, a competição da pecuária e as doenças trazidas pelos carneiros domésticos dizimaram os carneiros-das-montanhas. Atualmente, existem menos de 70.000 destes animais, dispersos pela América do Norte em populações fragmentadas e isoladas.

Em 1922, os biólogos da vida selvagem estabeleceram uma população destes carneiros no National Bison Range, uma extensão de terra de 18.000 acres entre as montanhas do noroeste de Montana. Naquele ano, 12 carneiros-das-montanhas, quatro machos (carneiros) e oito fêmeas (ovelhas), foram pegos no Banff National Park no Canadá e transportados para o National Bison Range. Não foram introduzidos outros animais nessa população nos 60 anos seguintes.

A princípio, a população de carneiros no National Bison Range cresceu, protegida da caça e da pecuária. Após 8 anos, a população tinha crescido para 90 carneiros, mas então começou a reduzir lentamente. O tamanho da população aumentava e diminuía com o passar dos anos, mas o número de carneiros caiu para 50 em 1985, e a população corria perigo. A variação genética era baixa em comparação com outras populações nativas de carneiros. A taxa de reprodução de machos e fêmeas caiu e o tamanho e a sobrevida eram menores do que nas populações mais saudáveis. A população do National Bison Range estava sofrendo com a deriva genética, uma força evolutiva que atua nas pequenas populações e provoca mudanças aleatórias no *pool* de genes e a perda da variação genética.

Para contrabalançar os efeitos negativos da deriva genética, os biólogos adicionaram cinco carneiros de outros rebanhos em Montana e Wyoming em 1985, imitando os efeitos da migração natural entre os rebanhos. Outros 10 carneiros foram introduzidos entre 1990 e 1994. Este fluxo de novos genes teve um efeito drástico na saúde genética da população. A variação genética dos carneiros aumentou de forma significativa. Os carneiros não consanguíneos (com novos genes) eram mais dominantes, mais suscetíveis para copular e mais prováveis de produzir descendentes. As ovelhas não consanguíneas tinham mais de duas vezes o sucesso reprodutivo anual das fêmeas endogâmicas. A sobrevida dos adultos aumentou após a introdução dos novos genes e lentamente, a população cresceu, alcançando 69 carneiros em 2003.

O carneiro-das-montanhas no National Bison Range ilustra um princípio importante da genética: populações pequenas perdem a variação genética com o passar do tempo em decorrência de deriva genética, com consequências catastróficas para a sobrevida e a reprodução. A introdução de nova variação genética em uma população endogâmica, chamada **resgate genético**, com frequência melhora drasticamente a saúde da população e pode garantir melhor sua sobrevida no longo prazo. Estes efeitos têm implicações importantes para o manejo da vida selvagem, assim como para como os organismos evoluem no mundo natural.

Este capítulo introduz a *genética de populações*, o ramo da genética que estuda a composição genética de grupos de indivíduos e como a composição genética de um grupo muda com o tempo. Os geneticistas populacionais voltam sua atenção em uma **população mendeliana**, que é um grupo de indivíduos endogâmicos, sexualmente reprodutivos que têm um conjunto de genes comum, o ***pool* de genes (**ou ***pool* genético)**. Uma população evolui com as mudanças no seu *pool* de genes, portanto, a genética de populações também é o estudo da evolução. Os geneticistas populacionais estudam a variação nos alelos dentro e entre grupos e as forças evolutivas responsáveis pela modelagem dos padrões da variação genética encontrada na natureza.

Neste capítulo, vamos aprender como o *pool* de genes de uma população é determinado e quais fatores são responsáveis por sua modelagem.

25.1 As frequências genotípicas e alélicas são usadas para descrever o *pool* de genes de uma população

Uma característica clara e penetrante da vida é a variabilidade. Estudantes em um colégio típico apresentam variação na cor dos olhos, cor do cabelo, pigmentação da pele, altura, peso, características faciais, tipo sanguíneo e suscetibilidade para numerosas doenças e distúrbios. É improvável que dois estudantes na classe tenham a mesma aparência.

Os humanos não são únicos na sua variabilidade extensiva (**Figura 25.1 A**), quase todos os organismos exibem variação no fenótipo. Por exemplo, as joaninhas apresentam alta variação nos seus padrões de pigmentação (**Figura 25.1 B**), os camundongos variam no tamanho do corpo, as cobras têm números diferentes de listras na sua couraça e as plantas têm diferente suscetibilidade a pragas. Boa parte da variação fenotípica é hereditária. O reconhecimento da magnitude da variação fenotípica fez Charles Darwin ter a ideia da evolução em decorrência de seleção natural. A variação genética é a base de toda evolução e a magnitude da variação genética em uma população afeta seu potencial para se adaptar à mudança ambiental.

Figura 25.1 Todos os organismos exibem variação genética. A. Variação significativa nos seres humanos. **B.** Variação nos padrões de manchas nas joaninhas asiáticas (*Parte A*: Michael Dwyer/Alamy.)

Na verdade, existe mais variação genética nas populações do que é visível no fenótipo. Parte da variação existe em nível molecular devido, em parte, à redundância do código genético, que permite que diferentes códons especifiquem o mesmo aminoácido. Assim, dois membros de uma população podem produzir a mesma proteína mesmo se suas sequências de DNA forem diferentes. Sequências de DNA entre os genes e íntrons em genes não codificam proteínas; parte da variação nestas sequências também tem pouco efeito no fenótipo.

Uma ferramenta importante, mas frequentemente mal compreendida usada na genética de populações, é o modelo matemático. Vamos parar um minuto para analisar o que é um modelo e como ele pode ser usado. Um modelo matemático descreve um processo como uma equação. Os fatores que podem influenciar o processo são representados por variáveis na equação; a equação define a forma na qual as variáveis influenciam o processo. A maioria dos modelos são representações simplificadas de um processo, porque é impossível analisar simultaneamente todos os fatores que influenciam; alguns fatores devem ser ignorados para examinar os efeitos de outros. Primeiro, um modelo pode analisar apenas um ou alguns fatores, mas após seus efeitos serem compreendidos, o modelo pode ser melhorado com o acréscimo de mais detalhes. E importante, até um simples modelo pode ser a fonte de discernimento valioso sobre como um processo é influenciado por variáveis-chave.

Antes de explorarmos os processos evolutivos que modelam a variação genética, precisamos ser capazes de descrever a estrutura genética de uma população. A forma comum para descrever essa estrutura é enumerar os tipos e as frequências de genótipos e alelos em uma população.

Como calcular as frequências genotípicas

Uma frequência é simplesmente uma proporção ou uma porcentagem, em geral expressa como uma fração decimal. Por exemplo, se 20% dos alelos em um determinado *locus* em uma população são *A*, diríamos que a frequência do alelo *A* na população é 0,20. Para populações maiores, nas quais uma determinação de genes de todos os membros é impossível, é usada uma amostra da população e as frequências genotípicas e alélicas são calculadas para esta amostra (ver *Capítulo 24*). As

frequências genotípicas e alélicas da amostra são, então, usadas para representar o *pool* de genes da população.

Para calcular uma **frequência genotípica**, simplesmente somamos o número de indivíduos com o genótipo e dividimos pelo número total de indivíduos na amostra (*N*). Para um *locus* com três genótipos *AA*, *Aa* e *aa*, a frequência (*f*) de cada genótipo é

$$f(AA) = \frac{\text{número de indivíduos } AA}{N} \qquad (25.1)$$

$$f(Aa) = \frac{\text{número de indivíduos } Aa}{N}$$

$$f(aa) = \frac{\text{número de indivíduos } aa}{N}$$

A soma de todas as frequências genotípicas é sempre igual a 1.

Como calcular as frequências alélicas

O *pool* de genes de uma população também pode ser descrito em termos das frequências alélicas. Sempre existem menos alelos do que genótipos, então o *pool* de genes de uma população pode ser descrito em menos termos quando as frequências alélicas são usadas. Em uma população com reprodução sexuada, os genótipos são apenas junções temporárias dos alelos. Como descrito pelo princípio da segregação de Mendel, os genótipos se desfazem a cada geração e os alelos individuais são transmitidos para a próxima geração pelos gametas. Assim, os tipos e números de alelos, em vez dos genótipos, têm uma real continuidade de uma geração para a seguinte e compõem o *pool* genético de uma população.

As frequências alélicas podem ser calculadas a partir do (1) número ou das (2) frequências dos genótipos. Para calcular a **frequência alélica** a partir dos números de genótipos, contamos o número de cópias de um determinado alelo encontrado em uma amostra e dividimos pelo número total de todos os alelos na amostra:

$$\text{Frequência de um alelo} = \frac{\text{número de cópias do alelo}}{\text{número de cópias de todos os alelos no } locus} \qquad (25.2)$$

Para um *locus* com apenas dois alelos (*A* e *a*), as frequências dos alelos são representadas pelos símbolos *p* e *q* e podem ser calculados como se segue:

$$p = f(A) = \frac{2n_{AA} + n_{Aa}}{2N} \qquad (25.3)$$

$$q = f(a) = \frac{2n_{aa} + n_{Aa}}{2N}$$

Em que n_{AA}, n_{Aa} e n_{aa} representam os números de *AA*, *Aa* e *aa* individuais e *N* representa o número total de indivíduos na amostra. Para obter o número de cópias do alelo no numerador da equação, adicionamos duas vezes o número de homozigotos (porque cada um tem duas cópias do alelo para o qual a frequência está sendo calculada) para o número de heterozigotos (porque cada um tem uma única cópia do alelo). Dividimos por 2*N* porque cada indivíduo diploide tem dois alelos em um *locus*. A soma das frequências alélicas é sempre igual a 1 ($p + q = 1$), então, após *p* ser obtida, *q* pode ser determinada pela subtração $q = 1 - p$.

Por outro lado, as frequências alélicas podem ser calculadas a partir das frequências genotípicas. Este cálculo é útil se as frequências genotípicas já estiverem calculadas e os números de genótipos diferentes não estiverem disponíveis. Para calcular uma frequência alélica a partir das frequências genotípicas, adicione a frequência do homozigoto para cada alelo para dividir a metade a frequência de cada heterozigoto (porque metade dos alelos do heterozigoto são de cada tipo):

$$p = f(A) = f(AA) + 1/2\, f(Aa)$$
$$q = f(a) = f(aa) + 1/2\, f(Aa) \qquad (25.4)$$

Obtemos os mesmos valores de p e q se calcularmos as frequências alélicas dos números de genótipos (Equação 25.3) ou a partir das frequências genotípicas (Equação 25.4). Um mesmo cálculo das frequências alélicas é fornecido no próximo Problema desenvolvido. **Resolva o Problema 16**

Loci com alelos múltiplos. Podemos usar os mesmos princípios para determinar as frequências dos alelos para *loci* com mais de dois alelos. Para calcular as frequências alélicas a partir dos números de genótipos, contamos o número de cópias de um alelo ao adicionar duas vezes o número de homozigotos ao número de heterozigotos que têm o alelo e dividir esta soma por duas vezes o número de indivíduos na amostra. Para um *locus* com três alelos (A^1, A^2 e A^3) e seis genótipos (A^1A^1, A^1A^2, A^2A^2, A^1A^3, A^2A^3 e A^3A^3), as frequências (p, q e r) dos alelos são

$$p = f(A^1) = \frac{2n_{A^1A^1} + n_{A^1A^2} + n_{A^1A^3}}{2N} \qquad (25.5)$$

$$q = f(A^2) = \frac{2n_{A^2A^2} + n_{A^1A^2} + n_{A^2A^3}}{2N}$$

$$r = f(A^3) = \frac{2n_{A^3A^3} + n_{A^1A^3} + n_{A^2A^3}}{2N}$$

Por outro lado, podemos calcular as frequências de alelos múltiplos a partir das frequências genotípicas ao estender a Equação 25.4. Mais uma vez, adicionamos as frequências do homozigoto à metade da frequência de cada genótipo heterozigoto que tem o alelo:

$$\begin{aligned} p &= f(A^1A^1) + 1/2\, f(A^1A^2) + 1/2\, f(A^1A^3) \\ q &= f(A^2A^2) + 1/2\, f(A^1A^2) + 1/2\, f(A^2A^3) \\ r &= f(A^3A^3) + 1/2\, f(A^1A^3) + 1/2\, f(A^2A^3) \end{aligned} \qquad (25.6)$$

Loci ligados ao X. Para calcular as frequências alélicas dos genes nos *loci* ligados ao X, aplicamos os mesmos princípios. Entretanto, devemos lembrar que uma fêmea tem dois cromossomos X e, portanto, tem dois alelos ligados ao X, enquanto um macho tem um único cromossomo X e um alelo ligado ao X.

Suponha que existam dois alelos em um *locus* ligado ao X, X^A e X^a. As fêmeas podem ser homozigotas (X^AX^A ou X^aX^a) ou heterozigotas (X^AX^a). Todos os machos são hemizigotos (X^AY ou X^aY). Para determinar a frequência do alelo X^A (p), primeiro, contamos o número de cópias de X^A: multiplicamos o número de fêmeas X^AX^A por dois e adicionamos o número de fêmeas X^AX^a e o número de machos X^AY. Dividimos a soma pelo número total de alelos no *locus*, que é duas vezes o número total de fêmeas mais o número de machos:

$$p = f(X^A) = \frac{2n_{X^AX^A} + n_{X^AX^a} + n_{X^AY}}{2n_{\text{fêmeas}} + n_{\text{machos}}} \qquad (25.7\text{ A})$$

De forma semelhante, a frequência do alelo X^a é

$$q = f(X^a) = \frac{2n_{X^aX^a} + n_{X^AX^a} + n_{X^aY}}{2n_{\text{fêmeas}} + n_{\text{machos}}} \qquad (25.7\text{ B})$$

As frequências dos alelos ligados ao X também podem ser calculadas a partir das frequências genotípicas ao adicionar a frequência das fêmeas homozigotas para o alelo, metade da frequência das fêmeas heterozigotas para o alelo e a frequência dos machos hemizigotos para o alelo:

$$\begin{aligned} p &= f(X^A) = f(X^AX^A) + 1/2\, f(X^AX^a) + f(X^AY) \\ q &= f(X^a) = f(X^aX^a) + 1/2\, f(X^AX^a) + f(X^aY) \end{aligned} \qquad (25.8)$$

Se você se lembrar da lógica por trás destes cálculos, você poderá determinar as frequências alélicas para qualquer conjunto de genótipos e não será necessário memorizar todas as fórmulas. **Resolva o Problema 18**

Conceitos

A genética de populações envolve a composição genética de uma população e como ela muda com o passar do tempo. O *pool* de genes de uma população pode ser descrito pelas frequências dos genótipos e alelos na população.

Checagem dos conceitos 1

Quais são algumas das vantagens de usar frequências alélicas para descrever o *pool* de genes de uma população em vez de usar frequências genotípicas?

Problema desenvolvido

Os antígenos MN do tipo sanguíneo humano são determinados por dois alelos codominantes, L^M e L^N (ver *Capítulo 5*). Os tipos sanguíneos NM e os genótipos correspondentes de 398 finlandeses de Karjala estão apresentados no quadro a seguir.

Fenótipo	Genótipo	Número
MM	L^ML^M	182
MN	L^ML^N	172
NN	L^NL^N	44

Fonte: W. C. Boyd, Genetics and the Races of Man (Boston: Little, Brown, 1950.)

Calcule as frequências genotípica e alélica no *locus* MN para a população de Karjala.

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

As frequências genotípica e alélica da população.

Quais informações são fornecidas para solucionar o problema?

Os números dos diferentes genótipos MN na amostra.

Etapas para a solução

As frequências genotípicas para a população são calculadas com a seguinte fórmula:

$$\text{Frequência genotípica} = \frac{\text{número de indivíduos com genótipo}}{\text{número total de indivíduos na amostra } (N)}$$

$$f(L^ML^M) = \frac{\text{número de indivíduos } L^ML^M}{N} = \frac{182}{398} = 0{,}457$$

$$f(L^ML^N) = \frac{\text{número de indivíduos } L^ML^N}{N} = \frac{172}{398} = 0{,}432$$

$$f(L^NL^N) = \frac{\text{número de indivíduos } L^NL^N}{N} = \frac{44}{398} = 0{,}111$$

As frequências alélicas podem ser calculadas a partir do número ou das frequências dos genótipos. Para calcular as frequências alélicas a partir dos números de genótipos, adicionamos o número de cópias do alelo e dividimos pelo número de cópias de todos os alelos naquele *locus*.

$$\text{Frequência de um alelo} = \frac{\text{número de cópias do alelo}}{\text{número de cópias de todos alelos}}$$

$$p = f(L^M) = \frac{(2n_{L^ML^M}) + (n_{L^ML^N})}{2N}$$

$$= \frac{2(182) + 172}{2(398)} = \frac{536}{796} = 0{,}673$$

$$q = f(L^N) = \frac{(2n_{L^NL^N}) + (n_{L^ML^N})}{2N}$$

$$= \frac{2(44) + 172}{2(398)} = \frac{260}{796} = 0{,}327$$

Para calcular as frequências alélicas a partir das frequências genotípicas, adicionamos a frequência do homozigoto para aquele genótipo à metade da frequência de cada heterozigoto que tenha aquele alelo:

$$p = f(L^M) = f(L^ML^M) + 1/2\ f(L^ML^N) = 0{,}457 + 1/2\ (0{,}432) = 0{,}673$$

$$q = f(L^N) = f(L^NL^N) + 1/2\ f(L^ML^N) = 0{,}111 + 1/2\ (0{,}432) = 0{,}327$$

Agora tente calcular as frequências genotípicas e alélicas ao resolver o **Problema 17** no final do capítulo.

25.2 A lei de Hardy-Weinberg descreve o efeito da reprodução nas frequências genotípicas e alélicas

O objetivo principal da genética de populações é compreender os processos que modelam o *pool* de genes de uma população. Primeiro, precisamos perguntar quais efeitos a reprodução e os princípios mendelianos têm sobre as frequências genotípicas e alélicas: Como a segregação dos alelos na formação do gameta e na combinação dos alelos na fertilização influenciam o *pool* de genes? A resposta a esta pergunta está na **lei de Hardy-Weinberg**, que está entre os princípios mais importantes da genética populacional.

A lei de Hardy-Weinberg foi formulada independentemente por G. H. Hardy e Wilhelm Weinberg em 1908 (foram alcançadas conclusões semelhantes por vários outros geneticistas ao mesmo tempo). A lei é, na verdade, um modelo matemático que avalia o efeito da reprodução sobre as frequências genotípicas e alélicas de uma população. Ela faz várias suposições simplificadas sobre a população e fornece duas previsões-chave se estas suposições forem cumpridas. Para um *locus* autossômico com dois alelos, a lei de Hardy-Weinberg pode ser afirmada como:

Suposições Se uma população for grande, com acasalamento aleatório, e não afetada por mutação, migração ou seleção natural então:
Previsão 1 As frequências alélicas de uma população não mudam e
Previsão 2 As frequências alélicas se estabilizam (não mudam) após uma geração nas proporções p^2 (a frequência de *AA*), $2pq$ (a frequência de *Aa*) e q^2 (a frequência de *aa*), em que p é igual à frequência do alelo *A* e q é igual à frequência do alelo *a*.

A lei de Hardy-Weinberg indica que, quando as suposições não são satisfeitas, a reprodução sozinha não altera as frequências alélicas ou genotípicas e as frequências alélicas determinam as frequências dos genótipos.

A afirmação de que as frequências genotípicas se estabilizam após uma geração significa que elas podem mudar na primeira geração após o acasalamento aleatório porque é necessária uma geração de acasalamento aleatório para produzir as proporções de Hardy-Weinberg dos genótipos. Subsequentemente, as frequências genotípicas, como as frequências alélicas, não se alteram desde que a população continue a satisfazer as suposições da lei de Hardy-Weinberg. Quando os genótipos estão nas proporções esperadas de p^2, $2pq$ e q^2, a população está no **equilíbrio de Hardy-Weinberg**.

Conceitos

A lei de Hardy-Weinberg descreve como a reprodução e os princípios mendelianos influenciam as frequências alélicas e genotípicas de uma população.

Checagem dos conceitos 2

Qual afirmação a seguir não é uma suposição da lei de Hardy-Weinberg?

a. As frequências alélicas (*p* e *q*) são iguais.
b. A população acasala aleatoriamente.
c. A população é grande.
d. A seleção natural não tem efeito.

Frequências genotípicas no equilíbrio de Hardy-Weinberg

Como as condições da lei de Hardy-Weinberg levaram a proporções genotípicas de p^2, $2pq$ e q^2? O princípio de segregação de Mendel diz que cada organismo tem dois alelos em um *locus* e que cada um dos dois alelos tem uma mesma probabilidade de passar para um gameta. Assim, as frequências dos alelos nos gametas serão as mesmas que as frequências dos alelos nos parentes. Suponha que temos uma população mendeliana na qual as frequências dos alelos *A* e *a* são *p* e *q*, respectivamente. Estas frequências também serão as mesmas nos gametas. Se o acasalamento for aleatório (uma das suposições da lei de Hardy-Weinberg), os gametas se reúnem em combinações aleatórias, que podem ser representadas por um quadrado de Punnett (**Figura 25.2**).

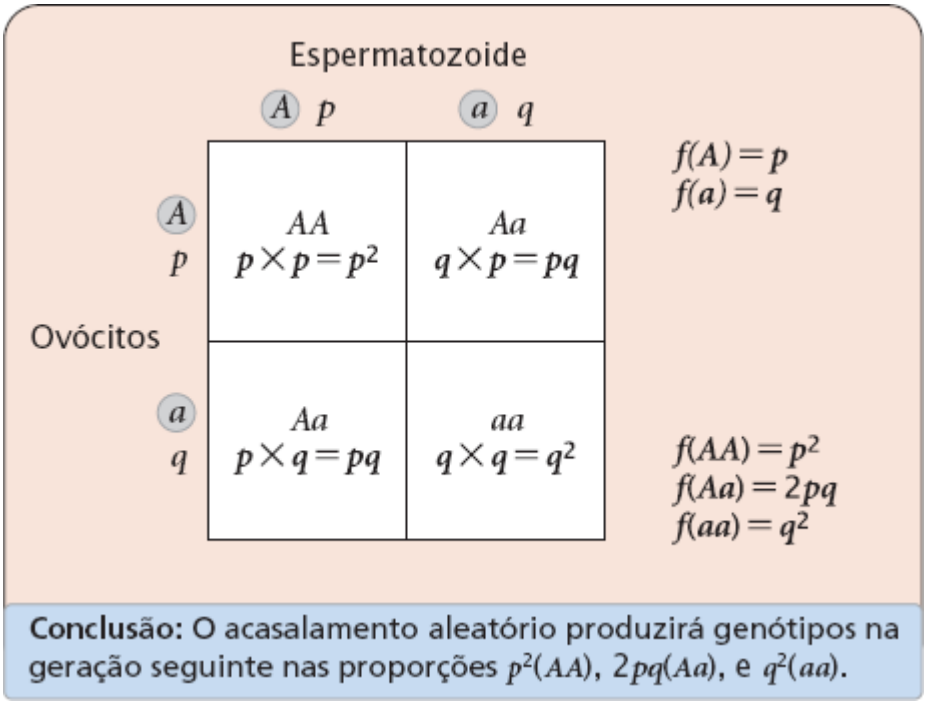


Figura 25.2 O acasalamento aleatório produz genótipos nas proporções p^2, $2pq$ e q^2.

A regra de multiplicação da probabilidade (ver *Capítulo 3*) pode ser usada para determinar a probabilidade de vários gametas parearem. Por exemplo, a probabilidade de um espermatozoide com o alelo *A* é *p* e a probabilidade de um ovócito com alelo *A* é *p*. Aplicando a regra da multiplicação, descobrimos que a probabilidade de estes dois gametas se combinarem para produzir um homozigoto *AA* é $p \times p = p^2$. Da mesma forma, a probabilidade de um espermatozoide com o alelo *a* combinar-se com um ovócito com o alelo *a* para produzir um homozigoto *aa* é $q \times q = q^2$. Um heterozigoto *Aa* pode ser produzido em uma das duas formas: (1) um espermatozoide com alelo *A* pode se combinar com um ovócito com o alelo *a* ($p \times q$) ou (2) um ovócito com alelo *A* pode se combinar com um espermatozoide contendo o alelo *a* ($p \times q$). Assim, a probabilidade de os alelos *A* e *a* se combinarem para produzir um heterozigoto *Aa* é $2pq$. Em resumo, sempre que as frequências dos alelos em uma população de acasalamento aleatório forem *p* e *q*, as frequências dos genótipos na próxima geração serão p^2, $2pq$ e q^2. A Figura 25.2 demonstra que apenas uma única geração de acasalamento aleatório é necessária para produzir as proporções genotípicas de Hardy-Weinberg.

Exame mais detalhado da lei de Hardy-Weinberg

Antes de analisar as implicações da lei de Hardy-Weinberg, precisamos examinar com mais detalhes as três suposições que ela faz sobre uma população. Primeiro, ela considera que a população é grande. Como quantificar o termo “grande”? Teoricamente, a lei de Hardy-Weinberg demanda que uma população seja infinitamente grande, mas esta exigência é obviamente irreal. Na prática, muitas populações grandes estão nas proporções previstas de Hardy-Weinberg e desvios significativos surgem apenas quando o tamanho da população é preferencialmente pequeno. Adiante no capítulo, vamos examinar os efeitos do tamanho pequeno de população sobre as frequências alélicas.

A segunda suposição da lei de Hardy-Weinberg é que os membros da população acasalam aleatoriamente, o que significa que cada genótipo acasala em relação a sua frequência. Por exemplo, suponha que existem três genótipos em uma população nas seguintes proporções: $f(AA) = 0,6$, $f(Aa) = 0,3$ e $f(aa) = 0,1$. Com o acasalamento aleatório, a frequência do acasalamento entre dois homozigotos AA (*AA* × *AA*) será igual à multiplicação de suas frequências: 0,6 × 0,6 = 0,36, enquanto a frequência de acasalamento entre dois homozigotos *aa* (*aa* × *aa*) será apenas 0,1 × 0,1 = 0,01.

A terceira suposição da lei de Hardy-Weinberg é que as frequências alélicas da população não são afetadas pela seleção natural, migração nem mutação. Embora a mutação ocorra em todas as populações, sua taxa é tão baixa que tem pouco efeito a curto prazo sobre as previsões desta lei (embora ela possa modelar em muito as frequências alélicas por longos períodos de tempo quando outras forças não estiverem agindo). Embora a seleção natural e a migração sejam fatores significativos nas populações reais, devemos lembrar que o objetivo da lei de Hardy-Weinberg é examinar apenas o efeito da reprodução sobre o *pool* de genes. Quando este efeito é conhecido, os efeitos de outros fatores (como a migração e a seleção natural) podem ser examinados.

Um ponto final é que as suposições da lei de Hardy-Weinberg se aplicam a um único *locus*. Nenhuma população real acasala aleatoriamente para todos os traços, e uma população não está completamente livre da seleção natural para todos os traços. A lei de Hardy-Weinberg, entretanto, não demanda acasalamento aleatório e a ausência de seleção, migração e mutação para todos os traços; ela requer que estas condições apenas para o *locus* em consideração. Uma população pode estar no equilíbrio de Hardy-Weinberg para um *locus*, mas não para outros.

Implicações da lei de Hardy-Weinberg

A lei de Hardy-Weinberg tem várias implicações importantes para a estrutura genética de uma população. Uma implicação é que uma população não pode evoluir se satisfizer as suposições de Hardy-Weinberg, porque a evolução consiste em alteração das frequências alélicas de uma população. Portanto, esta lei nos diz que a reprodução sozinha não produz evolução. São necessários outros processos, como a seleção natural, mutação, migração ou acaso para que as populações evoluam.

Uma segunda implicação importante é que, quando uma população está no equilíbrio de Hardy-Weinberg, as frequências genotípicas são determinadas pelas frequências alélicas. Quando uma população não está no equilíbrio de Hardy-Weinberg, não temos bases para prever as frequências genotípicas. Embora possamos sempre determinar as frequências alélicas a partir das frequências genotípicas (ver Equação 25.5), o inverso (determinar as frequências genotípicas a partir das frequências alélicas) é possível apenas quando a população está neste equilíbrio.

Para um *locus* com dois alelos, a frequência do heterozigoto é maior quando as frequências alélicas estão entre 0,33 e 0,66 e está no máximo quando as frequências alélicas são 0,5 cada (**Figura 25.3**). A frequência do heterozigoto também nunca excede 0,5 quando a população está no equilíbrio Hardy-Weinberg. Além disso, quando a frequência de um alelo está baixa, os homozigotos para aquele alelo serão raros, e a maioria das cópias de um alelo raro estará presente nos heterozigotos. Como você pode ver na Figura 25.3, quando a frequência do alelo *a* é 0,2, a frequência do homozigoto *aa* é apenas 0,04 (q^2), mas a frequência dos heterozigotos *Aa* é 0,32 ($2pq$); 80% dos alelos *a* estão nos heterozigotos.

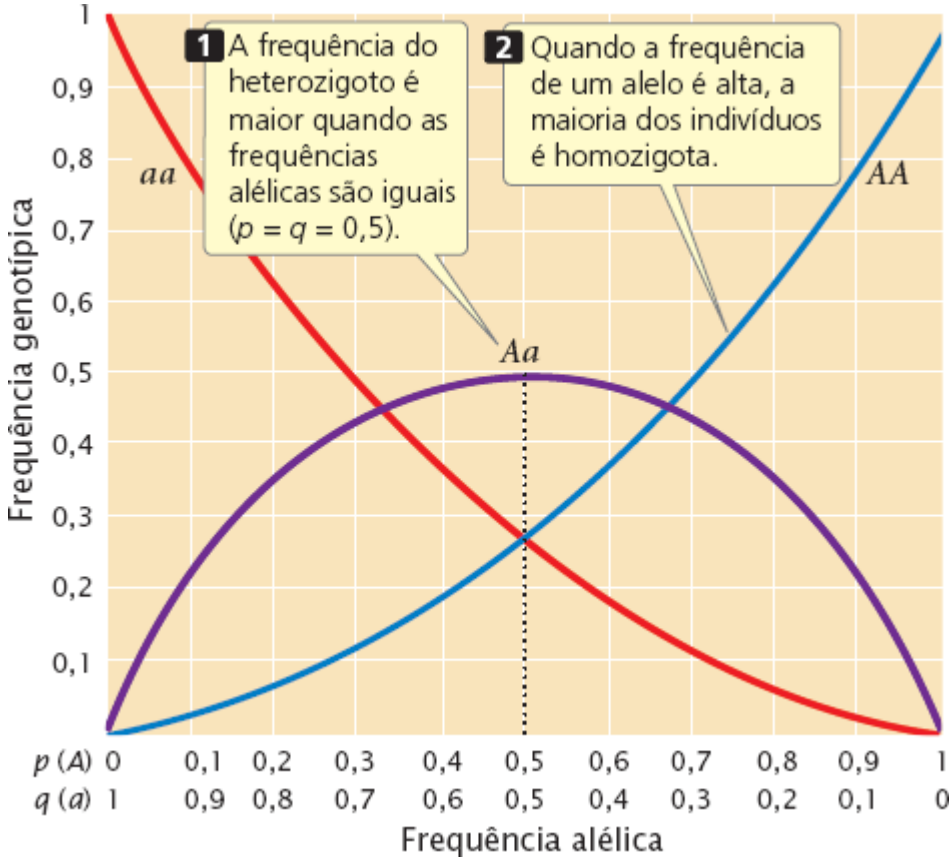


Figura 25.3 Quando uma população está no equilíbrio de Hardy-Weinberg, as proporções de genótipos são determinadas pelas frequências dos alelos.

Uma terceira implicação da lei de Hardy-Weinberg é que uma única geração de acasalamento aleatório produz as frequências de equilíbrio de p^2, $2pq$ e q^2. O fato de que os genótipos estão nas proporções de Hardy-Weinberg não prova que a população está livre de seleção natural, mutação e migração. Ela significa apenas que estas forças não atuam desde que ocorreu o último acasalamento aleatório.

Extensões da lei de Hardy-Weinberg

As proporções esperadas de Hardy-Weinberg também podem ser aplicadas a alelos múltiplos e alelos ligados ao X (**Quadro 25.1**). Com alelos múltiplos, as frequências genotípicas esperadas no equilíbrio são o quadrado das frequências alélicas. Para um *locus* autossômico com três alelos, as frequências genotípicas no equilíbrio serão $(p + q + r)^2 = p^2 + 2pq + q^2 + 2pr + 2qr + r^2$. Para um *locus* ligado ao X com dois alelos, X^A e X^a, as frequências de equilíbrio dos genótipos das fêmeas são $(p + q)^2 = p^2 + 2pq + q^2$, em que p^2 é a frequência de X^AX^A, $2pq$ é a frequência de X^AX^a e q^2 é a frequência de X^aX^a. Os machos têm apenas um único alelo ligado ao X, então as frequências dos genótipos dos machos são p (frequência de X^AY) e q (frequência de X^aY). Estas proporções são as dos genótipos entre os machos e fêmeas, em vez das proporções entre a população inteira. Assim, p^2 é a proporção esperada de fêmeas com genótipo X^AX^A; se as fêmeas compõem 50% da população, então a proporção esperada deste genótipo na população inteira é $0,5 \times p^2$.

A frequência de um traço recessivo ligado ao X entre os machos é q, enquanto a frequência entre as fêmeas é q^2. Quando um alelo ligado ao X é incomum, o traço será muito mais frequente nos machos que nas fêmeas. Considere a hemofilia A um distúrbio de coagulação causado por um alelo recessivo ligado ao X com uma frequência (q) de aproximadamente 1 em 10.000 ou 0,0001. No equilíbrio de Hardy-Weinberg, esta frequência também será a frequência da doença entre os homens. A frequência da doença entre as mulheres, entretanto, será $q^2 = (0,0001)^2 = 0,00000001$, que é apenas 1 em 100 milhões. A hemofilia é 10.000 vezes mais frequente nos homens que nas mulheres.

Quadro 25.1 Extensões da lei de Hardy-Weinberg.

Situação	Frequências alélicas	Frequências genotípicas
Três alelos	$f(A^1) = p$ $f(A^2) = q$ $f(A^3) = r$	$f(A^1A^1) = p^2$ $f(A^1A^2) = 2pq$ $f(A^2A^2) = q^2$

		$f(A^1A^3) = 2pr$ $f(A^2A^3) = 2qr$ $f(A^3A^3) = r^2$
Alelos ligados ao X	$f(X^1) = p$ $f(X^2) = q$	$f(X^1X^1$ fêmea$) = p^2$ $f(X^1X^2$ fêmea$) = 2pq$ $f(X^2X^2$ fêmea$) = q^2$ $f(X^1$Y macho$) = p$ $f(X^2$Y macho$) = q$

Nota: Para os genótipos de fêmeas ligados ao X, as frequências são as proporções entre todas as fêmeas; para os genótipos de machos ligados ao X, as frequências são as proporções entre todos os machos.

Teste para proporções de Hardy-Weinberg

Para determinar se os genótipos de uma população estão no equilíbrio de Hardy-Weinberg, as proporções genotípicas esperadas dentro da lei de Hardy-Weinberg devem ser comparadas com as frequências genotípicas observadas. Para tal, primeiro calculamos as frequências alélicas, então encontramos as frequências genotípicas esperadas ao usar o quadrado das frequências alélicas e, finalmente, comparamos as frequências genotípicas observadas e esperadas ao usar o teste qui-quadrado.

Problema desenvolvido

Jeffrey Mitton e colaboradores encontraram três genótipos (R^2R^2, R^2R^3 e R^3R^3) em um *locus* que codifica a enzima peroxidase em pinheiros que cresceram em Glacier Lake, Colorado. Os números observados destes genótipos foram:

Genótipos	Número observado
R^2R^2	135
R^2R^3	44
R^3R^3	11

Os pinheiros do *Glacier Lake* estão no equilíbrio de Hardy-Weinberg no *locus* da peroxidase?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

Os resultados de um teste qui-quadrado para determinar a população estão no equilíbrio de Hardy-Weinberg.

Quais informações são fornecidas para solucionar o problema?

Os números de diferentes genótipos em uma amostra da população.

Etapas para a solução

Se a frequência do alelo R^2 é igual a p e a frequência do alelo R^3 é igual a q, a frequência do alelo R^2 é

$$p = f(R^2) = \frac{(2n_{R^2R^2}) + (n_{R^2R^3})}{2N} = \frac{2(135) + 44}{2(190)} = 0{,}826$$

A frequência do alelo R^3 é obtida pela subtração:

$$q = f(R^3) = 1 - p = 0{,}174$$

As frequências dos genótipos esperados sob o equilíbrio de Hardy-Weinberg são calculadas ao usar p^2, $2pq$ e q^2:

$R^2R^2 = p^2 = (0{,}826)^2 = 0{,}683$

$R^2R^3 = 2pq = 2(0{,}826)(0{,}174) = 0{,}287$

$R^3R^3 = (0{,}174)^2 = 0{,}03$

Ao multiplicar cada uma dessas frequências genotípicas esperadas pelo número total de genótipos observados na amostra (190), obtemos os números esperados para cada genótipo:

$R^2R^2 = 0{,}683 \times 190 = 129{,}8$

$R^2R^3 = 0{,}287 \times 190 = 54{,}5$

$R^3R^3 = 0{,}03 \times 190 = 5{,}7$

Ao comparar estes números esperados com os números observados de cada genótipo, observamos que existem mais homozigotos R^2R^2 e R^3R^3 e menos heterozigotos R^2R^3 na população do que esperamos no equilíbrio.
Um teste qui-quadrado de adequação do ajuste é usado para determinar se as diferenças entre os números observados e esperados de cada genótipo são devido ao acaso:

$$X^2 = \sum \frac{(\text{observado} - \text{esperado})^2}{\text{esperado}}$$

$$= \frac{(135 - 129{,}8)^2}{129{,}8} + \frac{(44 - 54{,}5)^2}{54{,}5} + \frac{(11 - 5{,}7)^2}{5{,}7}$$

$$= 0{,}21 + 2{,}02 + 4{,}93 = 7{,}16$$

O valor calculado de qui-quadrado é 7,16; para obter a probabilidade associada com este valor qui-quadrado, determinamos os adequados graus de liberdade.

Até este ponto, o teste qui-quadrado para avaliar o equilíbrio de Hardy-Weinberg é idêntico aos testes qui-quadrado que usamos no Capítulo 3 para avaliar as razões de descendentes em um cruzamento genético, em que os graus de liberdade eram $n - 1$ e n igual ao número de genótipos esperados. Para o teste Hardy-Weinberg, entretanto, devemos subtrair um grau adicional de liberdade, como os números esperados baseiam-se nas frequências alélicas observadas; portanto, os números observados não estão totalmente livres de variar. Em geral, os graus de liberdade para um teste qui-quadrado do equilíbrio de Hardy-Weinberg são iguais ao número de classes genotípicas esperadas menos o número de alelos associados. Para esse teste de Hardy-Weinberg particular, o grau de liberdade é 3 – 2 = 1.

Após calcularmos o valor de qui-quadrado e os graus de liberdade, a probabilidade associada com este valor pode ser obtida em um quadro de qui-quadrado (ver Quadro 3.7). Com um grau de liberdade, um valor de qui-quadrado de 7,16 tem uma probabilidade entre 0,1 e 0,001. É muito improvável que os genótipos de peroxidase observados no Glacier Lake estejam nas proporções de Hardy-Weinberg.

Para exercitar mais, determine se as frequências genotípicas no **Problema 22** no final do capítulo estão no equilíbrio de Hardy-Weinberg.

Conceitos

O número observado de genótipos em uma população pode ser comparado com as proporções esperadas de Hardy-Weinberg ao usar um teste qui-quadrado de adequação de ajuste.

Checagem dos conceitos 3

Qual é a frequência esperada de heterozigotos em uma população com frequências alélicas x e y que está no equilíbrio de Hardy-Weinberg?

a. $x + y$.
b. xy.
c. $2xy$.
d. $(x - y)^2$.

Como estimar as frequências alélicas com a lei de Hardy-Weinberg

Um uso prático da lei de Hardy-Weinberg é que ela permite que calculemos as frequências alélicas quando a dominância está presente. Por exemplo, a fibrose cística é um distúrbio autossômico recessivo caracterizado por infecções respiratórias, digestão incompleta e transpiração anormal (ver *Capítulo 5*). Entre os caucasianos norte-americanos, a incidência da doença é de aproximadamente 1 pessoa em 2.000. A fórmula para calcular a frequência alélica (ver Equação 25.3) requer que saibamos os números de homozigotos e heterozigotos, mas a fibrose cística é uma doença recessiva, então não podemos distinguir entre pessoas homozigotas não afetadas e os carreadores heterozigotos. Embora estejam disponíveis testes moleculares para identificar os carreadores heterozigotos do gene da fibrose cística, a baixa frequência da doença torna a triagem difundida pouco prática. Em tais casos, a lei de Hardy-Weinberg pode ser usada para estimar as frequências alélicas.

Se considerarmos que uma população está no equilíbrio de Hardy-Weinberg com relação a este *locus*, então a frequência do genótipo recessivo (*aa*) será q^2 e a frequência alélica é a raiz quadrada da frequência genotípica:

$$q = \sqrt{f(aa)} \tag{25.9}$$

Se a frequência da fibrose cística nos caucasianos norte-americanos é aproximadamente 1 em 2.000 ou 0,0005, então $q = \sqrt{0{,}0005}$ = 0,02. Assim, cerca de 2% dos alelos na população caucasiana codificam a fibrose cística. Podemos calcular a frequência do alelo normal ao subtrair $p = 1 - q = 1 - 0{,}02 = 0{,}98$. Após calcularmos *p* e *q*, podemos usar a lei de Hardy-Weinberg para determinar as frequências das pessoas homozigotas não afetadas e os carreadores heterozigotos do gene:

$$f(AA) = p^2 = (0{,}98)^2 = 0{,}960$$
$$f(Aa) = 2pq = 2(0{,}02)(0{,}98) = 0{,}0392$$

Assim, cerca de 4% (1 de 25) dos caucasianos são carreadores heterozigotos do alelo que provoca fibrose cística. **Resolva o Problema 25**

Conceitos

Embora as frequências alélicas não possam ser calculadas diretamente para os traços que exibem dominância, a lei de Hardy-Weinberg pode ser usada para estimar as frequências alélicas se a população estiver no equilíbrio de Hardy-Weinberg para aquele *locus*. A frequência do alelo recessivo será igual à raiz quadrada da frequência do traço recessivo.

Checagem dos conceitos 4

Nos gatos, a cor branca é dominante sobre a não branca. Em uma população de 100 gatos, 19 são todos brancos. Assumindo que a população está no equilíbrio de Hardy-Weinberg, qual é a frequência do alelo todo branco nesta população?

25.3 O acasalamento não aleatório afeta as frequências genotípicas de uma população

Uma suposição da lei de Hardy-Weinberg é que o acasalamento é aleatório com relação ao genótipo. A acasalamento não aleatório afeta a forma na qual os alelos se combinam para formar genótipos e altera as frequências genotípicas de uma população.

Podemos distinguir entre os dois tipos de acasalamento não aleatório. O **acasalamento preferencial positivo** se refere a uma tendência para indivíduos próximos em acasalar. Por exemplo, os humanos exibem acasalamento preferencial positivo para altura: pessoas altas acasalam preferencialmente com outras pessoas altas; pessoas baixas acasalam preferencialmente com outras pessoas baixas. O **acasalamento preferencial negativo** se refere a uma tendência para indivíduos distantes em acasalar. Se as pessoas são envolvidas no acasalamento preferencial negativo pela altura, as pessoas altas e baixas vão acasalar preferencialmente. O acasalamento preferencial é usado em um traço particular e afetará apenas os genes que codificam o traço (e genes intimamente ligados a este).

Uma forma de acasalamento não aleatório é a **endogamia**, que é o acasalamento preferencial entre indivíduos próximos. A

endogamia é na verdade o acasalamento preferencial positivo para o parentesco, mas difere dos outros tipos de acasalamento preferencial porque afeta todos os genes, não apenas os que determinam o traço para o qual existe a preferência de acasalamento. A endogamia causa uma divergência nas frequências de equilíbrio de Hardy-Weinberg de p^2, $2pq$ e q^2. Mais especificamente, ela leva a um aumento na proporção de homozigotos e uma redução na proporção de heterozigotos em uma população. O **exocruzamento** ou **exogamia** (***outcrossing***) é evitar o acasalamento entre indivíduos próximos.

Em um organismo diploide, um indivíduo homozigoto tem duas cópias do mesmo alelo. Estas duas cópias podem ser a mesma por *estado*, o que significa que os dois alelos são semelhantes em estrutura e função, mas não têm uma descendência comum. Por outro lado, os dois alelos em um indivíduo homozigoto podem ser o mesmo porque são idênticos por *descendência*; ou seja, as cópias são descendentes de um único alelo que estava presente em um ancestral (**Figura 25.4**). Se voltarmos muito no tempo, é provável que muitos alelos sejam idênticos por descendência, mas para calcular os efeitos da endogamia consideramos a identidade por descendência ao retornarmos apenas algumas gerações.

A endogamia é medida pelo **coeficiente de endogamia** designado como F, que é uma medida da probabilidade de que dois alelos sejam idênticos por descendência. Os coeficientes de endogamia podem variar de 0 a 1. Um valor de 0 indica que o acasalamento em uma população grande é aleatório; um valor de 1 indica que todos os alelos são idênticos por descendência. Os coeficientes de endogamia podem ser calculados a partir de análises de heredogramas ou podem ser determinados pela redução da heterozigosidade de uma população. Não vamos entrar em detalhes sobre como F é calculado, mas é importante compreender com a endogamia afeta as frequências genotípicas.

Quando ocorre a endogamia, a proporção de heterozigotos reduz-se para $2Fpq$, e metade deste valor (Fpq) é adicionado à proporção de cada homozigoto em cada geração. As frequências dos genótipos serão então

$$\begin{aligned} f(AA) &= p^2 + Fpq \\ f(Aa) &= 2pq + 2Fpq \\ f(aa) &= q^2 + Fpq \end{aligned} \qquad (25.10)$$

Considere uma população que se reproduz por autofertilização ($F = 1$). Vamos considerar que esta população começa com frequências genotípicas nas proporções de Hardy-Weinberg (p_2, $2pq$ e q_2). Com a autofecundação, cada homozigoto produz descendentes apenas do mesmo genótipo homozigoto ($AA \times AA$ produz todos AA e $aa \times aa$ produz todos aa), enquanto apenas metade dos descendentes de um heterozigoto será como o genitor ($Aa \times Aa$ produz 1/4 AA 1/2 Aa e 1/4 aa). A autofecundação, portanto, reduz a proporção de heterozigotos na população a metade com cada geração, até todos os genótipos na população serem homozigotos (**Quadro 25.2** e **Figura 25.5**). Embora todos os indivíduos na população sejam homozigotos neste ponto, ambos tipos de homozigotos (AA e aa) ainda estão presentes, desde que nenhum homozigoto tenha vantagem sobre o outro (não existem diferenças seletivas). **Resolva o Problema 28**

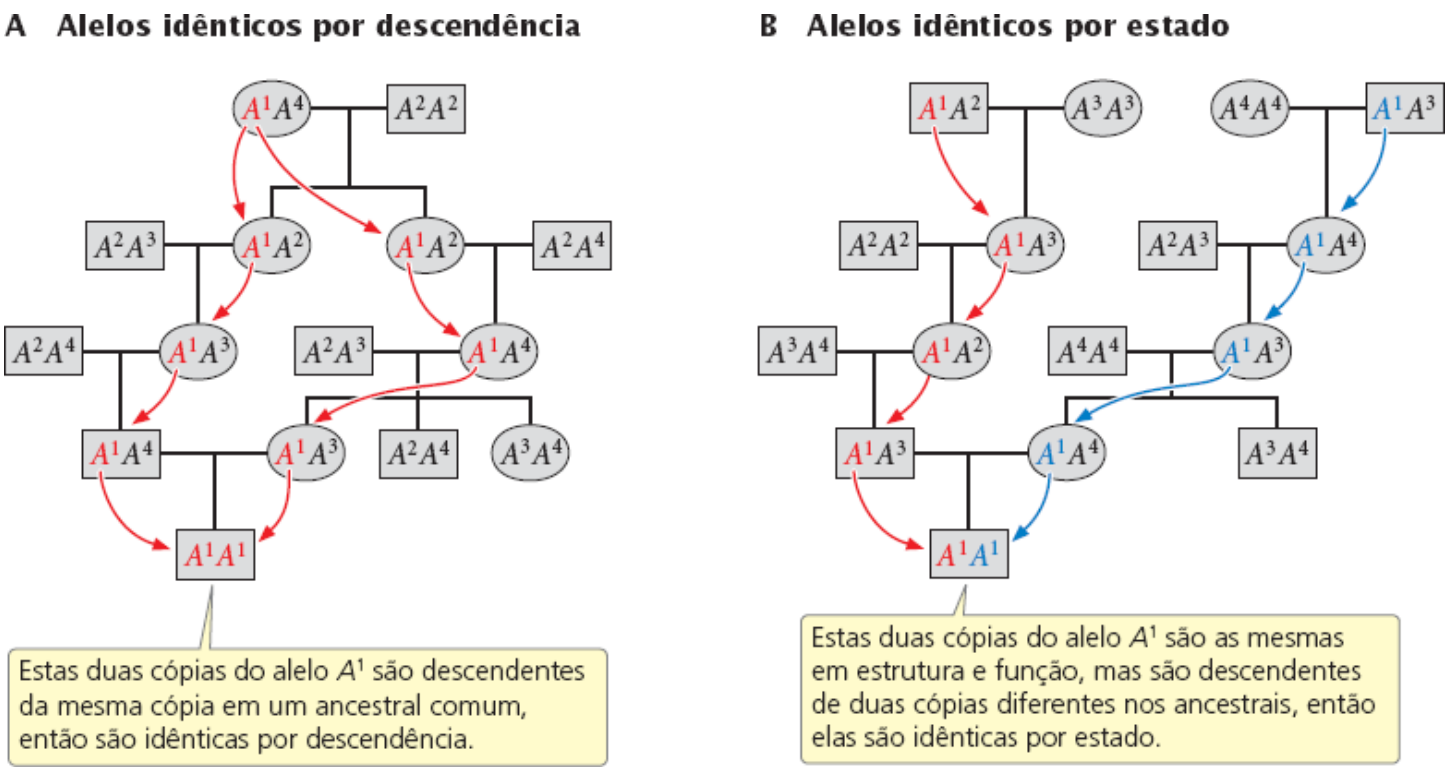


Figura 25.4 Os indivíduos podem ser homozigotos por descendência ou por estado. A endogamia é uma medida da

probabilidade de que dois alelos sejam idênticos na origem.

Para a maioria das espécies de cruzamento exogâmico, a endogamia próxima é prejudicial porque aumenta a proporção de homozigotos e, portanto, aumenta a probabilidade de alelos recessivos deletérios ou letais se combinarem para produzir homozigotos com um traço prejudicial. Considere que um alelo recessivo (*a*) que causa uma doença genética tem uma frequência (*q*) 0,01. Se a população se acasala aleatoriamente ($F = 0$), a frequência dos indivíduos afetados com a doença (*aa*) será $q^2 = 0{,}01^2 = 0{,}0001$, então apenas 1 em 10.000 indivíduos terá a doença. Entretanto, se $F = 0{,}25$ (equivalente a acasalamento entre irmãos), então a frequência esperada do genótipo homozigoto é $q^2 + Fpq = (0{,}01)^2 + (0{,}25)(0{,}99)(0{,}01) = 0{,}0026$; assim a doença genética é 26 vezes mais frequente neste nível de cruzamento. Este maior surgimento de traços letais e deletérios com a endogamia é chamado de **depressão da endogamia**; quanto mais intensa a endogamia, mais grave a depressão da endogamia.

Quadro 25.2 Aumento por geração na frequência de homozigotos em uma população se autofertilizando.*

Geração	Frequências genotípicas		
	AA	***Aa***	***aa***
1	1/4	1/2	1/4
2	1/4 + 1/8 = 3/8	1/4	1/4 + 1/8 = 3/8
3	3/8 + 1/16 = 7/16	1/8	3/8 + 1/16 = 7/16
4	7/16 + 1/32 = 15/32	1/16	7/16 + 1/32 = 15/32
n	$\frac{1-(1/2)^n}{2}$	$(1/2)^n$	$\frac{1-(1/2)^n}{2}$
	1/2	0	1/2

*Começando com $p = q = 0{,}5$.

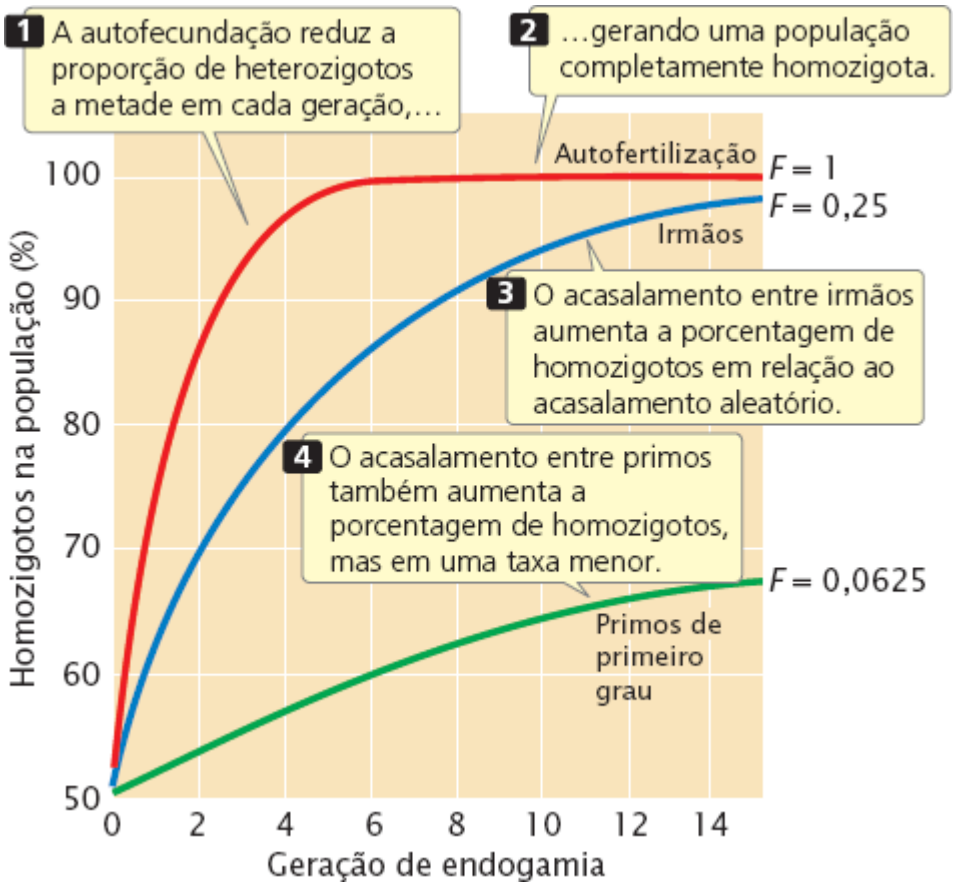


Figura 25.5 A endogamia aumenta a porcentagem de indivíduos homozigotos em uma população.

Os efeitos prejudiciais da endogamia são conhecidos há milênios e podem ser a base do tabu cultural proibindo o acasalamento entre parentes. William Schull e James Neel descobriram que cada aumento de 10% em *F*, a média do QI das crianças japonesas caía seis pontos. A mortalidade infantil também aumenta com a endogamia (**Quadro 25.3**); os filhos de primos de primeiro grau têm um aumento de 40% na mortalidade sobre a mortalidade entre os filhos de pessoas não consanguíneas. A endogamia também tem efeitos prejudiciais nos cultivos (**Figura 25.6**) e animais domésticos.

A depressão da endogamia é mais estudada nos humanos, assim como nas plantas e animais criados em cativeiro, mas os efeitos negativos da endogamia podem ser mais intensos nas populações naturais. Julie Jimenez e colaboradores coletaram camundongos selvagens de uma população natural em Illinois e os cruzaram no laboratório por três ou quatro gerações. Os acasalamentos do laboratório foram escolhidos de modo que alguns camundongos não tivessem endogamia, enquanto outros tiveram um coeficiente de endogamia de 0,25. Quando ambos tipos de camundongos foram liberados de volta para o *habitat* selvagem, a sobrevida semanal dos camundongos de endogamia foi de apenas 56% daquela dos camundongos que não tiveram este cruzamento. Os camundongos machos intercruzados perderam o peso corporal continuamente após serem liberados para o *habitat*, enquanto os camundongos não intercruzados inicialmente perderam peso, mas então ganharam dentro de alguns dias após a liberação.

Quadro 25.3 Efeitos da endogamia nas crianças japonesas.

Relação genética dos genitores	*F*	Mortalidade das crianças (até 12 anos)
Não consanguíneo	0	0,082
Primos de segundo grau	0,016 (1/64)	0,108
Primos de primeiro grau	0,0625 (1/16)	0,114

Fonte: De D. L. Hartl and A. G. Clark, *Principles of Population Genetics*, 2d ed. (Sunderland, Mass.: Sinauer, 1989), Table 2. Original data from W. J. Schull and J. V. Neel, *The Effects of Inbreeding on Japanese Children* (New York: Harper & Row, 1965).

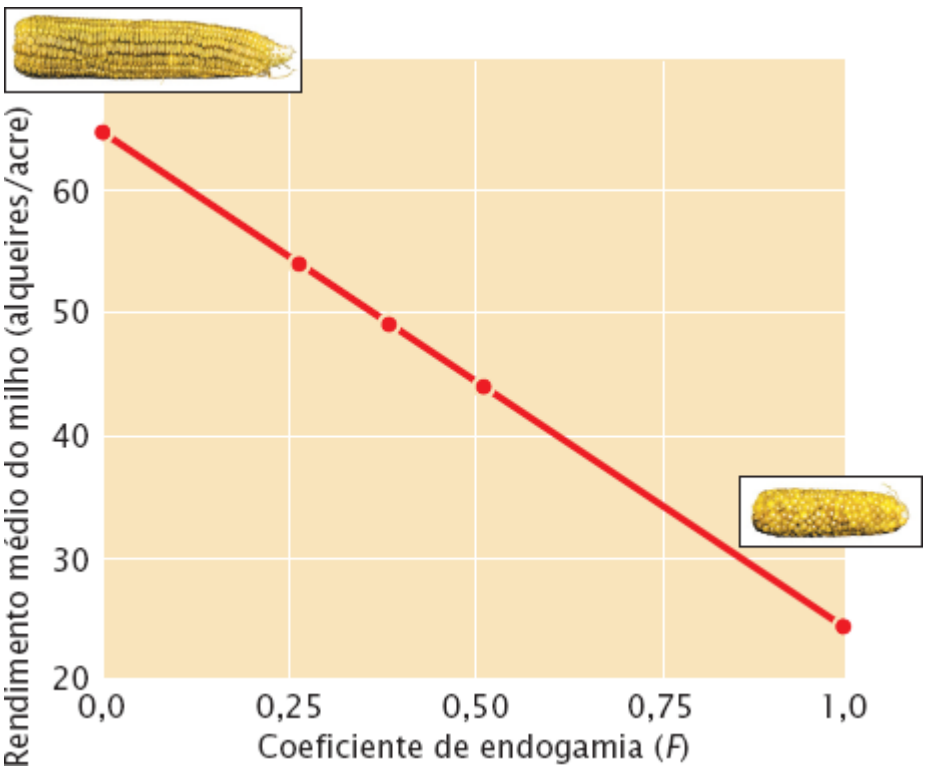


Figura 25.6 A endogamia tem efeitos prejudiciais nos cultivos. À medida que a endogamia aumenta, o rendimento médio do milho, por exemplo, reduz-se.

Apesar de a endogamia ser prejudicial para espécies exogâmicas, várias plantas e animais intercruzam regularmente e com sucesso. Como foi afirmado anteriormente, a endogamia aumenta a homozigosidade e consequentemente todos os indivíduos na população se tornam homozigotos. Se uma espécie sofre endogamia por várias gerações, muitos alelos recessivos prejudiciais

são extirpados pela seleção natural ou artificial de modo que a população se torna homozigota para alelos benéficos. Desta forma, os efeitos prejudiciais da endogamia podem ser eliminados, deixando uma população homozigota para os traços benéficos.

Conceitos

O acasalamento não aleatório altera as frequências dos genótipos, mas não as frequências dos alelos. A endogamia é o acasalamento preferencial entre indivíduos próximos. Com a endogamia, a frequência de homozigotos aumenta, enquanto a frequência de heterozigotos diminui.

Checagem dos conceitos 5

Qual o efeito do exocruzamento em uma população?

a. As frequências alélicas mudam.
b. Existem mais heterozigotos do que o previsto pela lei de Hardy-Weinberg.
c. Existem menos heterozigotos do que o previsto pela lei de Hardy-Weinberg.
d. As frequências genotípicas serão iguais às previstas pela lei de Hardy-Weinberg.

25.4 Várias forças evolutivas causam mudanças nas frequências alélicas

A lei de Hardy-Weinberg indica que as frequências alélicas não mudam como resultado da reprodução. Os processos que produzem mudança na frequência alélica incluem mutação, migração, deriva genética (efeitos aleatórios por causa de tamanho pequeno da população) e seleção natural.

Mutação

Antes que a evolução possa ocorrer, deve existir variação genética dentro de uma população; consequentemente, toda evolução depende de processos que geram a variação genética. Embora possam surgir novas *combinações* de genes existentes pela recombinação na meiose, todos as variantes genéticas surgem pela mutação.

Efeito da mutação nas frequências alélicas. A mutação pode influenciar a taxa na qual uma variante genética aumenta à custa de outra. Considere um único *locus* em uma população de 25 indivíduos diploides. Cada indivíduo tem dois alelos no *locus* em análise, então o *pool* de genes da população tem 50 cópias de alelos. Vamos assumir que existam dois alelos diferentes, designados G^1 e G^2 com as frequências p e q, respectivamente. Se existem 45 cópias de G^1 e 5 cópias de G^2 na população, $p = 0{,}90$ e $q = 0{,}10$. Agora, suponha que ocorra uma mutação em um alelo G^1 para um alelo G^2. Após esta mutação, existem 44 cópias de G^1 e 6 cópias de G^2 e a frequência de G^2 aumentou de 0,10 para 0,12. A mutação mudou a frequência alélica.

Se as cópias de G^1 continuarem a sofrer mutação para G^2, a frequência de G^2 vai aumentar e a frequência de G^1 vai diminuir (**Figura 25.7**). A quantidade que G^2 vai mudar (Δq) como resultado da mutação depende de: (1) a taxa da mutação G^1 para G^2 (μ) e (2) p, a frequência de G^1 na população. Quando p é grande, existem muitas cópias de G^1 disponíveis para sofrer mutação para G^2, e a mudança será relativamente grande. À medida que ocorrem mais mutações e p diminui, existirão menos cópias de G^1 disponíveis para sofrer mutação para G^2. A mudança em G^2 como resultado da mutação é igual à taxa de mutação multiplicada pela frequência alélica:

$$\Delta q = \mu p \tag{25.11}$$

À medida que a frequência de p reduz-se como resultado da mutação, a mudança na frequência por causa da mutação será menor e menor.

Até o momento, consideramos apenas os efeitos das mutações diretas $G^1 \rightarrow G^2$. As mutações reversas $G^2 \rightarrow G^1$ também ocorrem na taxa v, que provavelmente será diferente da taxa de mutação direta, μ. Sempre que ocorre uma mutação reversa, a frequência de G^2 reduz-se e a frequência de G^1 aumenta (ver Figura 25.7). A taxa de mudança por causa das mutações reversas é igual à taxa de mutação reversa vezes a frequência alélica de G^2 ($\Delta q = vq$). A mudança geral na frequência alélica é um equilíbrio entre as forças opostas da mutação direta e da mutação reversa:

$$\Delta q = \mu p - v p \tag{25.12}$$

Como alcançar o equilíbrio das frequências alélicas. Considere uma população que começa com uma alta frequência de G^1 e baixa frequência de G^2. Nesta população, muitas cópias de G^1 estão inicialmente disponíveis para sofrer mutação para G^2 e o aumento em G^2 pela mutação direta será relativamente grande. Entretanto, à medida que a frequência de G^2 aumenta como resultado das mutações diretas, menos cópias de G^1 estão disponíveis para mutação, então o número de mutações diretas diminui; por outro lado, poucas cópias de G^2 estão disponíveis para sofrer uma mutação reversa para G^1, mas à medida que a frequência de G^2 aumenta, o número de cópias de G^2 disponíveis para sofrer mutação reversa para G^1 aumenta, então o número de genes sofrendo mutação reversa vai aumentar. Consequentemente, o número de genes que sofrem mutação direta será contrabalançado pelo número de genes que sofrem mutação reversa. Neste ponto, o aumento em q por causa da mutação direta será igual à redução em q por causa da mutação reversa e não existirá mudança real na frequência alélica ($\Delta q = 0$), apesar do fato de que as mutações direta e reversa continuem a ocorrer. O ponto no qual não existe mudança na frequência alélica de uma população é chamado de **equilíbrio** (ver Figura 25.7). No equilíbrio, a frequência de G^2 $\hat{q}$ será

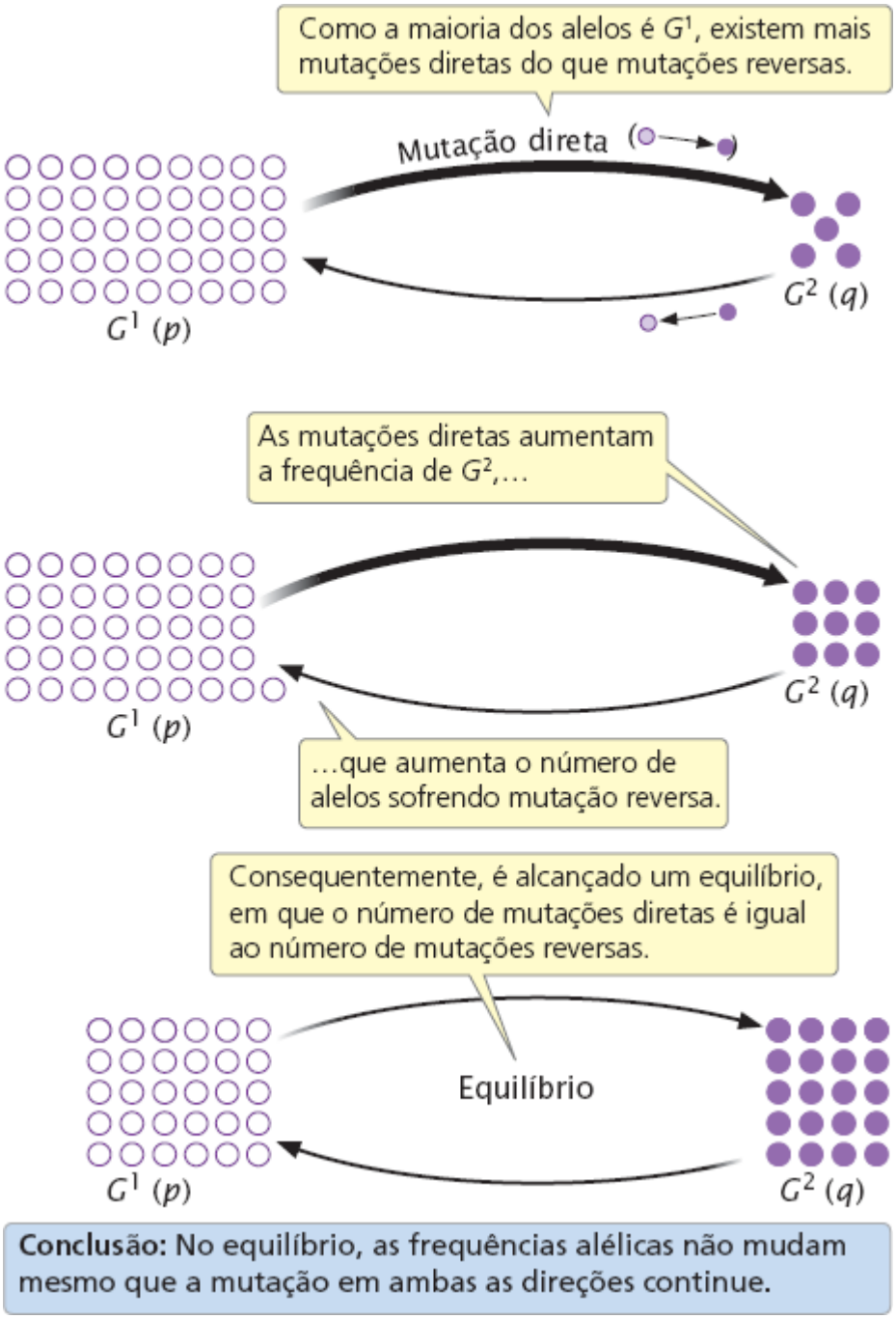


Figura 25.7 A mutação recorrente muda as frequências alélicas. As mutações direta e reversa levam ao equilíbrio.

$$\hat{q} = \frac{\mu}{\mu + \nu} \tag{25.13}$$

Esta equação final nos diz que a frequência alélica no equilíbrio é determinada apenas pelas taxas de mutação direta (μ) e

reversa (v). **Resolva o Problema 30**

Resumo dos efeitos. Quando a única força evolutiva que atua em uma população é a mutação, as frequências alélicas mudam com o passar do tempo porque alguns alelos sofrem mutação para outros. Consequentemente, estas frequências alélicas alcançam o equilíbrio e são determinadas apenas pelas taxas de mutação direta e reversa. Quando as frequências alélicas alcançam o equilíbrio, a lei de Hardy-Weinberg nos diz que as frequências genotípicas também permanecem as mesmas.

As taxas de mutação para a maioria dos genes são baixas, então mudanças na frequência alélica devido a mutação em uma geração são muito pequenas, e são necessários longos períodos de tempo para que uma população alcance um equilíbrio mutacional. Por exemplo, se as taxas de mutação direta e reversa para alelos em um *locus* são 1×10^{-5} e $0{,}3 \times 10^{-5}$ por geração, respectivamente (as taxas que são medidas em vários *loci* em camundongos), e as frequências alélicas são $p = 0{,}9$ e $q = 0{,}1$, então a mudança total na frequência alélica por geração por causa de mutação é

$$\begin{aligned}\Delta q &= \mu p - v q \\ &= (1 \times 10^{-5})(0{,}9) - (0{,}3 \times 10^{-5})(0{,}1) \\ &= 8{,}7 \times 10^{-6} = 0{,}0000087\end{aligned}$$

Portanto, a mudança causada por mutação em uma única geração é muito pequena e como a frequência de p cai como resultado da mutação, a quantidade de mutação será ainda menor (**Figura 25.8**). O efeito das taxas típicas de mutação no equilíbrio de Hardy-Weinberg é desprezível, e são necessárias muitas gerações para uma população alcançar o equilíbrio mutacional. Contudo, se a mutação é a única força que atua em uma população por muito tempo, as taxas de mutação determinarão as frequências alélicas.

Conceitos

Mutações recorrentes provocam mudanças nas frequências dos alelos. No equilíbrio, as frequências alélicas são determinadas pelas taxas de mutação direta e reversa. Como as taxas de mutação são baixas, o efeito da mutação nas frequências alélicas por geração é muito pequeno.

Checagem dos conceitos 6

Quando uma população está em equilíbrio para as taxas de mutação direta e reversa, qual das afirmações a seguir é verdadeira?

a. O número de mutações diretas é maior que o número de mutações reversas.
b. Não ocorre mutação direta nem mutação reversa.
c. O número de mutações diretas é igual ao número de mutações reversas.
d. A população está no equilíbrio Hardy-Weinberg.

Migração

Outro processo que pode provocar mudança nas frequências alélicas é a entrada de genes de outras populações, comumente chamada de **migração** ou **fluxo gênico**. Uma das suposições da lei de Hardy-Weinberg é que não ocorre migração, mas muitas populações naturais experimentam a migração de outras populações. O efeito geral da migração é duplo: (1) evita que populações se tornem geneticamente diferentes uma da outra e (2) aumenta a variação genética dentro das populações.

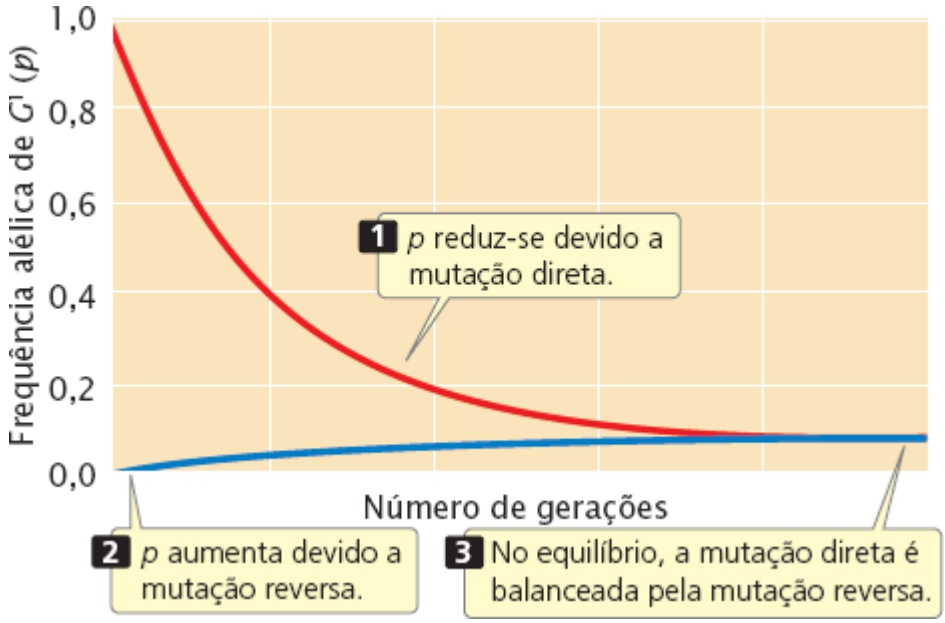


Figura 25.8 Mudança devido a mutação recorrente reduz-se à medida que a frequência de *p* cai. As frequências alélicas aproximam-se do equilíbrio mutacional em taxas de mutações lentas. A frequência alélica de G^1 reduz como resultado da mutação direta ($G^1 \rightarrow G^2$) em uma taxa de 0,0001 e aumenta como resultado da mutação reversa ($G^2 \rightarrow G^1$) em uma taxa de 0,00001. Devido à baixa taxa de mutações, são necessárias muitas gerações para alcançar o equilíbrio.

Efeito da migração nas frequências alélicas. Vamos analisar os efeitos da migração ao observar um modelo unidirecional único de migração entre duas populações com diferença na frequência de um alelo *a*. Considere que a frequência deste alelo na população I é q_I e que na população II é q_{II} (**Figura 25.9**). Em cada geração, uma amostra representativa dos indivíduos na população I migra para a população II e se reproduz, adicionando seus alelos ao *pool* de genes da população II. A migração é apenas da população I para a população II (unidirecional) e todas as condições da lei de Hardy-Weinberg (população de tamanho grande, acasalamento aleatório etc.) aplicam-se, exceto a ausência de migração.

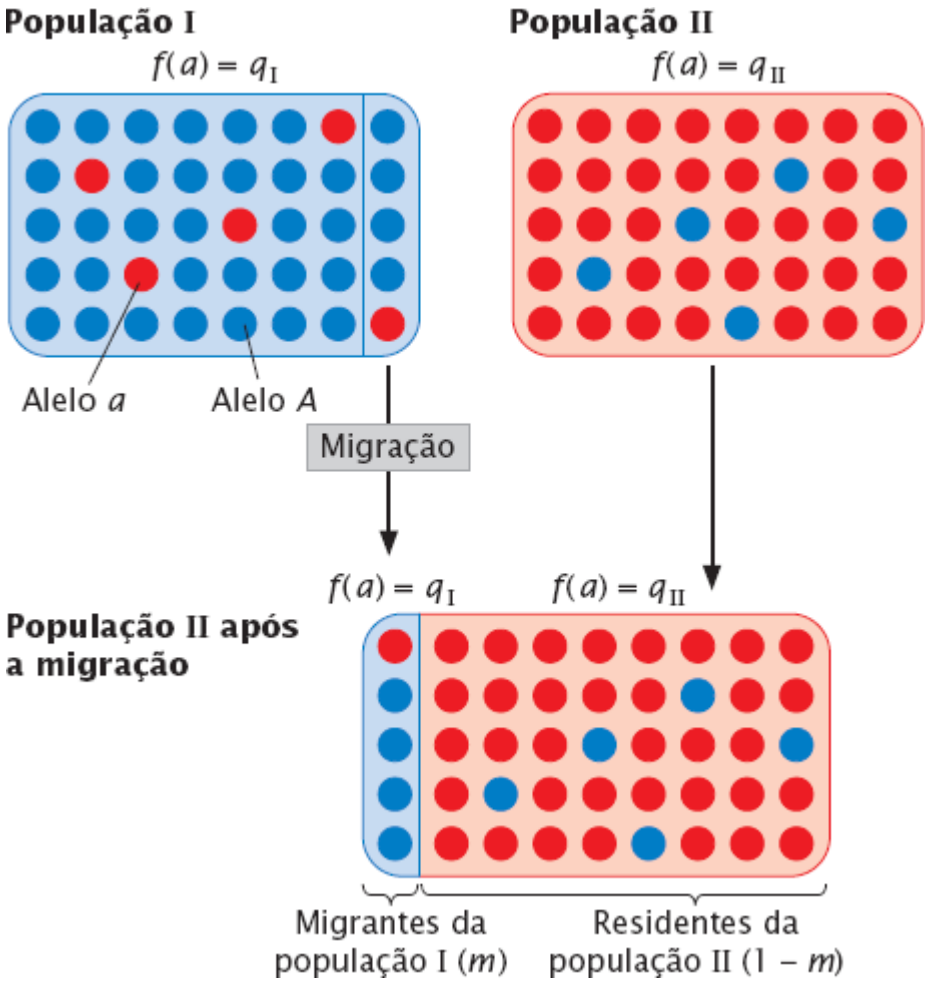


Figura 25.9 A quantidade de mudança na frequência alélica devido à migração entre populações depende da diferença na frequência alélica e na extensão da migração. Aqui está representado um modelo do efeito da migração unidirecional sobre as frequências alélicas. A frequência do alelo *a* na população de origem (população I) é q_I. A frequência deste alelo na população receptora (população II) é q_{II}.

Após a migração, a população II é composta por dois tipos de indivíduos. Alguns são migrantes; eles compõem a proporção m da população II e carreiam alelos da população I, então a frequência do alelo *a* nos migrantes é q_I. Outros indivíduos na população II são os residentes originais. Se os imigrantes compõem a proporção m da população II, então os residentes compõem $1 - m$; como os residentes se originaram na população II, a frequência do alelo *a* neste grupo é q_{II}. Após a migração, a frequência do alelo *a* na população II misturada (q'_{II}) é

$$q'_{II} = q_I\,(m) + q_{II}\,(1 - m) \tag{25.14}$$

Em que $q_I(m)$ é a contribuição para q feita por cópias do alelo *a* nos migrantes e $q_{II}(1 - m)$ é a contribuição para q feita por cópias do alelo *a* nos residentes. A mudança na frequência alélica devido a mutação (Δq) será

$$\Delta q = m(qI_I - q_{II}) \tag{25.15}$$

A Equação 25.15 resume os fatores que determinam a quantidade de mudança na frequência alélica devido a mutação. A quantidade de mudança em q é diretamente proporcional à migração (m); à medida que aumenta a migração a mudança na frequência alélica aumenta. A magnitude da mudança também é afetada pelas diferenças nas frequências alélicas das duas populações ($q_I - q_{II}$); quando a diferença é grande, a mudança na frequência alélica será grande.

Com cada geração de migração, as frequências das duas populações se tornam cada vez mais semelhantes até, consequentemente, a frequência alélica da população II ser igual à da população I. Quando $q_I - q_{II} = 0$, não existirá mais diferença na frequência alélica da população II, apesar de a migração continuar. Se a migração entre duas populações ocorrer por várias gerações, sem outras forças evolutivas presentes, é alcançado um equilíbrio no qual a frequência alélica da população receptora é igual à da população de origem.

O modelo único de migração unidirecional entre duas populações destacado aqui também pode ser expandido para acomodar a migração multidirecional entre várias populações.

Efeito geral da migração. A migração tem dois efeitos principais. Primeiro, ela faz com os *pools* de genes das populações se tornem mais semelhantes. Posteriormente, veremos como a deriva genética e a seleção natural levam a diferenças genéticas entre as populações, a migração anula esta tendência e tende a manter as populações homogêneas nas suas frequências alélicas. Segundo, a migração adiciona variação genética às populações. Diferentes alelos podem surgir em diferentes populações devido a raros eventos de mutação e estes alelos podem ser difundidos a novas populações pela migração, aumentando a variação genética dentro da população receptora. Resolva o Problema 34

Conceitos

A migração provoca mudanças na frequência alélica de uma população ao introduzir alelos a partir de outras populações. A magnitude da mudança causada pela migração depende da extensão da migração e da diferença nas frequências alélicas entre as populações de origem e receptora. A migração diminui as diferenças genéticas entre as populações e aumenta a variação genética dentro das populações.

Checagem dos conceitos 7

Em cada geração, 10 indivíduos aleatórios migram da população A para a população B. O que acontecerá à frequência alélica q como resultado da migração quando q é igual nas populações A e B?

a. q em A vai diminuir.
b. q em B vai aumentar.
c. q não aumentará em A nem em B.
d. q em B se tornará q^2.

Deriva genética

A lei de Hardy-Weinberg assume o acasalamento aleatório em uma população infinitamente grande; apenas quando o tamanho da população é infinito os gametas carrearão genes que representam perfeitamente o *pool* de genes parental. Mas nenhuma

população é infinitamente grande e quando o tamanho da população é limitado, os gametas que se unem para formar os indivíduos da próxima geração carreiam uma amostra dos genes presentes no *pool* de genes parental. Apenas por acaso, a composição desta amostra se desviará do *pool* de genes parental e este desvio pode fazer com que as frequências alélicas mudem. Quanto menor a amostra genética, maior a chance de que sua composição se desvie da composição do *pool* de genes inteiro.

O papel do acaso em alterar as frequências alélicas é análogo ao de jogar uma moeda. Cada vez que jogamos uma moeda, temos 50% de chance de ter cara e 50% de chance de ter coroa. Se jogarmos uma moeda 1.000 vezes, a razão observada de caras para coroas será muito próxima da razão 50:50 esperada. Se, entretanto, jogarmos uma moeda apenas 10 vezes, existe uma boa chance de que teremos não exatamente cinco caras e cinco coroas, mas talvez sete caras e três coroas ou oito coroas e duas caras. Este tipo de desvio de uma razão esperada devido ao tamanho limitado da amostra é chamado de **erro de amostragem**.

O erro de amostragem surge quando os gametas se unem para produzir descendentes. Muitos organismos produzem vários gametas, mas quando a população é pequena, um número limitado de gametas se une para produzir os indivíduos da próxima geração. O acaso influencia quais alelos estão presentes nesta amostra limitada e, desta forma, o erro de amostragem pode levar à **deriva genética**, ou mudanças na frequência alélica. Como os desvios das razões esperadas são aleatórios, a direção do acaso é imprevisível. Contudo, não podemos nunca prever a magnitude das mudanças.

Magnitude da deriva genética. O efeito da deriva genética pode ser observado de duas formas. Primeiro, podemos ver como ela influencia a mudança nas frequências alélicas de uma única população com o passar do tempo. Segundo, podemos ver como ela afeta as diferenças que se acumulam entre várias populações. Imagine que temos 10 populações pequenas, todas começando com as mesmas frequências alélicas exatas $p = 0,5$ e $q = 0,5$. Quando ocorre a deriva genética em uma população, as frequências alélicas dentro da população vão mudar, mas como a deriva genética é aleatória, a forma na qual as frequências alélicas mudam em cada população não será a mesma. Em algumas populações, p pode aumentar como resultado do acaso. Em algumas populações, p pode diminuir como resultado do acaso. Com o tempo, as frequências alélicas nas 10 populações serão diferentes: as populações vão divergir do ponto de vista genético. À medida que o tempo passa, a mudança na frequência alélica dentro de uma população e a divergência genética entre as populações ocorre pela mesma força – a mudança aleatória nas frequências alélicas. A magnitude da deriva genética pode ser avaliada ao examinar a mudança na frequência alélica dentro de uma única população ou ao examinar a magnitude das diferenças genéticas que se acumulam entre as populações.

A quantidade de deriva genética pode ser estimada a partir da variância na frequência alélica. A variância, s^2, é uma medida estatística que descreve o grau de variabilidade em um traço (ver *Capítulo 24*). Suponha que observemos um grande número de populações separadas, cada uma com N indivíduos e frequências p e q. Após uma geração de acasalamento aleatório, a deriva genética expressa em termos de variância na frequência alélica entre as populações (s_p^2) será

$$s_p^2 = \frac{pq}{2N} \tag{25.16}$$

A mudança resultante da deriva genética (a variância na frequência alélica) é determinada por dois parâmetros: as frequências alélicas (p e q) e o tamanho da população (N). A deriva genética será máxima quando p e q forem iguais (cada um 0,5). Por exemplo, considere que uma população tem 50 indivíduos. Quando as frequências alélicas forem iguais ($p = q = 0,5$), a variância na frequência alélica (sp^2) será $(0,5 \times 0,5)/(2 \times 50) = 0,0025$. Por outro lado, quando $p = 0,9$ e $q = 0,1$, a variância na frequência alélica será apenas 0,0009. A deriva genética também será maior quando a população for pequena. Se $p = q = 0,5$, mas o tamanho da população é apenas 10, em vez de 50, então a variância na frequência alélica é $(0,5 \times 0,5)/(2 \times 10) = 0,0125$, que é cinco vezes maior que quando a população é 50.

Esta divergência de populações pela deriva genética é ilustrada de forma clara nos resultados de um experimento feito por Peter Buri nas moscas-da-fruta (**Figura 25.10**). Buri examinou as frequências de dois alelos (bw^{75} e bw) que afetam a cor dos olhos nestas moscas. Ele montou 107 populações, cada uma com oito machos e oito fêmeas. Ele começou cada população com uma frequência de bw^{75} igual a 0,5. Permitiu que as moscas dentro de cada grupo acasalassem de forma aleatória e de cada geração ele selecionava aleatoriamente oito machos e oito fêmeas para serem os genitores da próxima geração. Ele acompanhou as mudanças nas frequências dos dois alelos por 19 gerações. Em uma população, a frequência média de bw^{75} (p) sobre as 19 gerações foi 0,5312. Podemos usar a Equação 25.16 para calcular a variância esperada na frequência alélica causada pela deriva genética. A frequência do alelo bw (q) será $1 - p = 1 - 0,53125 = 0,46875$. O tamanho da população (N) é igual a 16. A variância esperada na frequência alélica será

$$\frac{pq}{2N} = \frac{0,53125 \times 0,46875}{2 \times 16} = 0,0156$$

que estava muito próxima da variância real observada de 0,0151.

O efeito do tamanho da população na deriva genética está ilustrado por um estudo conduzido por Luca Cavalli-Sforza e colaboradores. Eles estudaram a variação nos tipos sanguíneos entre aldeões no Vale de Parma, na Itália, onde a migração entre os vilarejos era limitada. Eles descobriram que a variação na frequência alélica era a maior entre as vilas pequenas isoladas no vale superior, mas reduzida entre as vilas maiores e cidades na parte inferior do vale. Este resultado era exatamente o que esperamos com a deriva genética: deve haver mais deriva genética e assim mais variação entre as vilas quando a população é pequena.

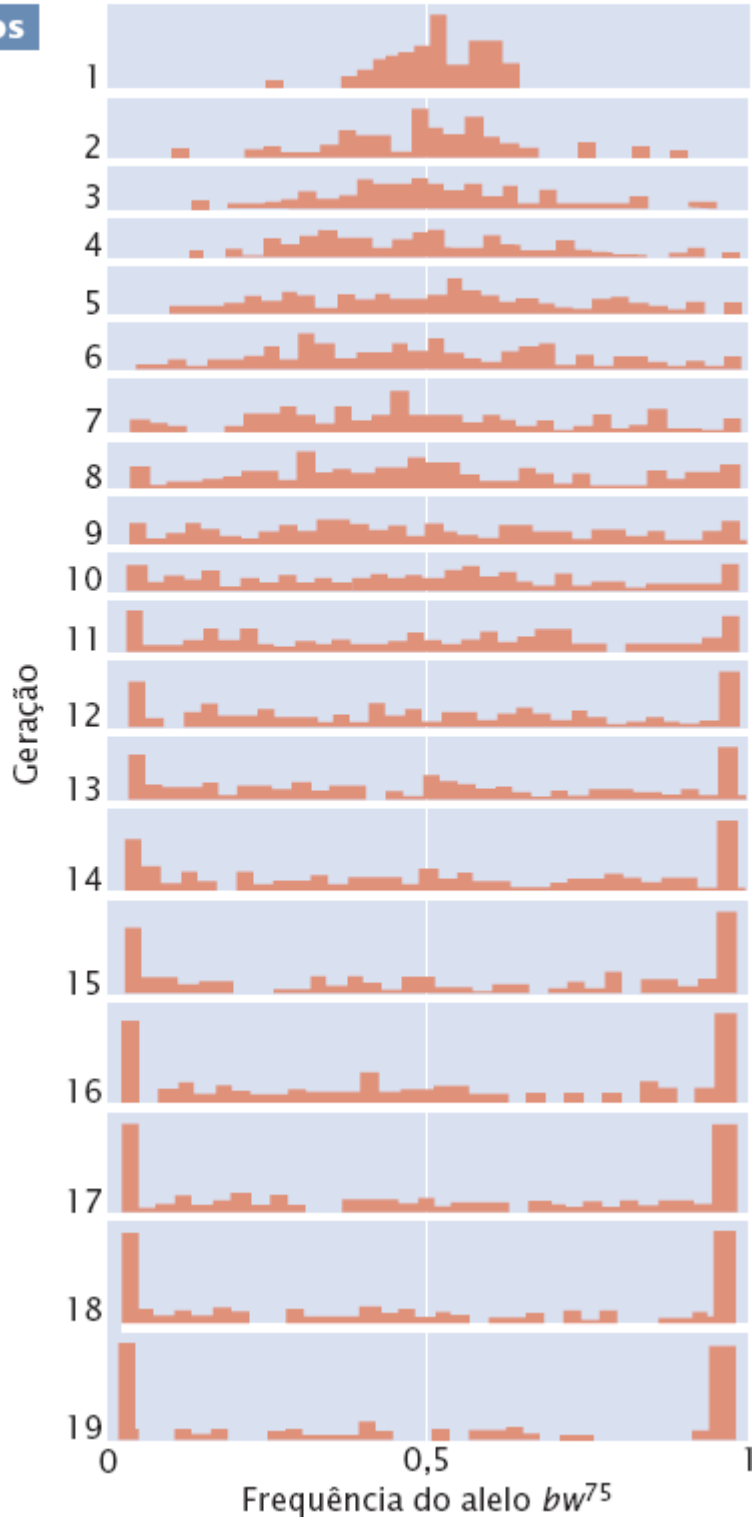
Experimento

Pergunta: Qual efeito a deriva genética tem sobre a composição genética das populações?

Métodos

Buri examinou as frequências de dois alelos (bw^{75} e *bw*) que afetam a cor dos olhos na *Drosophila* em 107 pequenas populações replicadas por 19 gerações.

Resultados



Conclusão: Como resultado da deriva genética, as frequências alélicas em diferentes populações divergiam e eram fixas para um alelo ou outro.

Figura 25.10 As populações divergem na frequência alélica e se tornam fixas para um alelo como resultado da deriva genética. No estudo de Bury dos dois alelos de cor dos olhos (bw^{75} e *bw*) na *Drosophila*, cada população consistia em oito machos e oito fêmeas e começou com a frequência de bw^{75} igual a 0,5.

Para estudos ecológicos e demográficos, o tamanho da população é definido como o número de indivíduos em um grupo. Entretanto, a evolução de um *pool* de genes depende apenas dos indivíduos que contribuem com genes para a próxima geração.

Os geneticistas de população em geral definem o tamanho da população como o número equivalente de adultos para acasalar, o **tamanho efetivo da população** (N_e). Vários fatores determinam o número equivalente de adultos para acasalar, incluindo razão de sexos, variação entre indivíduos no sucesso reprodutivo, flutuações no tamanho da população, estrutura etária da população e se o acasalamento é aleatório.

Conceitos

A deriva genética é a mudança na frequência alélica causada por fatores do acaso. A quantidade de mudança na frequência alélica causada pela deriva genética é inversamente o tamanho da população efetiva (o número equivalente de adultos reprodutores em uma população).

Checagem dos conceitos 8

Qual das afirmações a seguir é um exemplo de deriva genética?

a. O alelo *g* para produção de gordura aumenta em uma pequena população porque os pássaros com mais gordura corporal têm maior sobrevivência em um inverno rigoroso.
b. A mutação aleatória aumenta a frequência do alelo *A* em uma população, mas não em outra.
c. O alelo *R* alcança uma frequência de 1,0 porque os indivíduos com o genótipo *rr* são estéreis.
d. O alelo *m* é perdido quando um vírus mata alguns indivíduos e apenas por acaso nenhum dos sobreviventes tem o alelo *m*.

Causas da deriva genética. Toda deriva genética surge a partir do erro de amostragem, mas existem várias formas diferentes nas quais o erro de amostragem pode surgir. Primeiro, uma população pode ter o tamanho reduzido por várias gerações por causa de limitações de espaço, alimento ou outro recurso crítico. A deriva genética em uma pequena população por múltiplas gerações pode afetar de forma significativa a composição do *pool* de genes de uma população.

Uma segunda forma na qual o erro de amostragem pode surgir é pelo **efeito do fundador**, que é devido ao estabelecimento de uma população por um pequeno número de indivíduos; a população de carneiro-das-montanhas no National Bison Range, discutida na introdução deste capítulo, sofreu o efeito do fundador. Embora uma população possa aumentar e se tornar bem grande, os genes carreados por todos seus membros são derivados de alguns genes originalmente presentes nos fundadores (assumindo que não haja migração nem mutação). Os eventos de acaso que afetam quais genes estavam presentes nos fundadores terão importante influência na composição da população inteira.

Uma terceira forma na qual a deriva genética surge é pelo **gargalo genético**, que se desenvolve quando uma população sofre uma redução drástica no seu tamanho. Um exemplo é observado nos elefantes-marinhos do norte. Antes de 1800, milhares de elefantes-marinhos do norte eram encontrados pela costa da Califórnia, mas a caça entre 1820 e 1880 devastou a população. Em 1884, apenas cerca de 20 elefantes-marinhos sobreviveram em uma praia remota da Ilha de Guadalupe, a Oeste de Baja California, México. As restrições à caça promulgadas pelos EUA e México permitiram que os elefantes-marinhos se recuperassem e agora existem mais de 30.000 elefantes-marinhos na população. Todos os elefantes-marinhos na população são geneticamente semelhantes porque eles têm genes que foram carreados por alguns sobreviventes do gargalo da população.

Efeitos da deriva genética. A deriva genética tem várias implicações importantes para a composição genética de uma população. Primeiro, ela produz mudança nas frequências alélicas dentro de uma população. Como a deriva é aleatória, a probabilidade da frequência em aumentar é a mesma que de reduzir e vai ficar à deriva com o passar do tempo (por isso o nome deriva genética). A **Figura 25.11** ilustra uma simulação de computador da deriva genética em cinco populações por 30 gerações, começando com $q = 0,5$ e mantendo uma população constante de 10 machos e 10 fêmeas. Estas frequências alélicas mudam aleatoriamente de geração em geração.

Um segundo efeito da deriva genética é a redução na variação genética dentro das populações. Pela mudança aleatória, um alelo pode alcançar uma frequência de 1 ou 0, ponto no qual todos os indivíduos na população são homozigotos para um alelo. Quando um alelo alcançou uma frequência de 1, dizemos que ele alcançou a **fixação**. Outros alelos são perdidos (alcançam uma frequência 0) e podem ser restaurados apenas pela migração de outra população ou pela mutação. A fixação leva a perda da variação genética dentro de uma população. Esta perda pode ser observada nos elefantes-marinhos do norte. Atualmente, esses elefantes-marinhos têm baixos níveis de variação genética; um estudo de 24 genes que codificam proteínas não encontrou diferenças individuais ou populacionais nesses genes. Um estudo subsequente da variância de sequência no DNA mitocondrial também revelou baixos níveis de variação genética. Por outro lado, os elefantes-marinhos do sul tinham níveis muito maiores de variação do DNA mitocondrial. Os elefantes-marinhos do sul, que são encontrados em águas antárticas e subantárticas, também foram caçados, mas sua população nunca ficou menor que 1.000; portanto, diferente dos elefantes do norte, eles não tiveram um

gargalo genético.

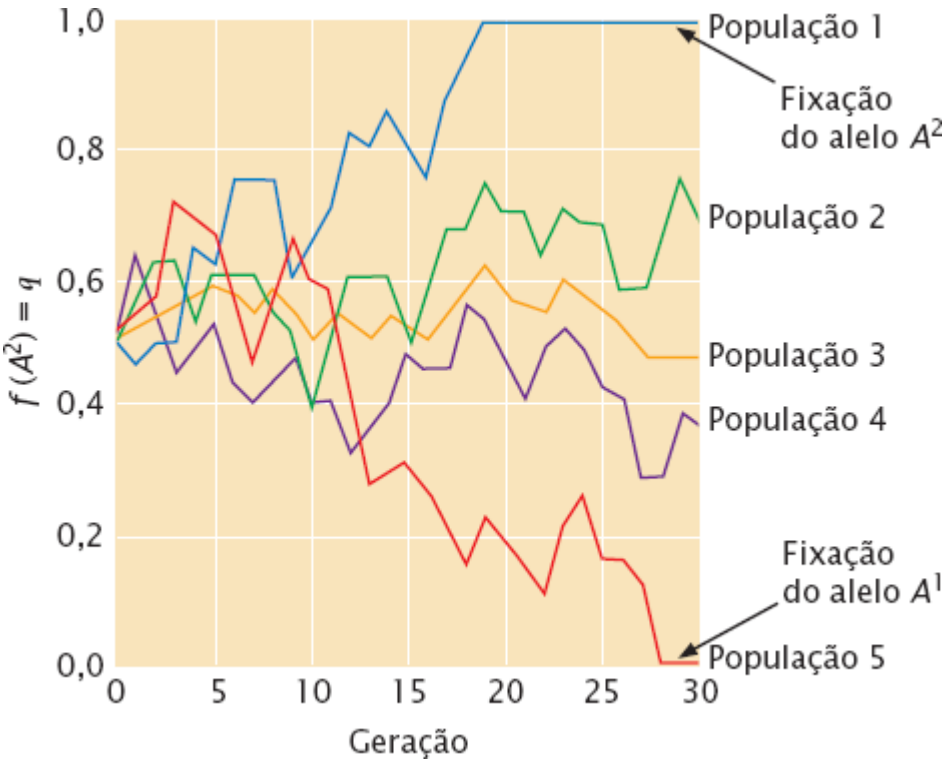


Figura 25.11 A deriva genética muda as frequências alélicas dentro das populações, levando a uma redução na variação genética pela fixação e divergência genética entre as populações. Aqui temos uma simulação de computador das mudanças na frequência do alelo A^2 (q) em cinco populações diferentes causadas pela deriva genética aleatória. Cada população é composta por 10 machos e 10 fêmeas e começa com q = 0,5.

Com tempo suficiente, todas as populações pequenas se tornarão fixas para um alelo ou outro. Qual alelo se torna fixo é aleatório e é determinado pela frequência inicial do alelo. Se uma população começa com dois alelos, cada um com uma frequência de 0,5, ambos alelos têm a mesma probabilidade de fixação. Entretanto, se um alelo é inicialmente comum, é mais provável que se torne fixo.

Um terceiro efeito da deriva genética é que diferentes populações divergem geneticamente com o tempo. Na Figura 25.11, todas as cinco populações começam com a mesma frequência alélica (q = 0,5) mas, como a deriva é aleatória, as frequências nas diferentes populações não mudam da mesma forma, e então as populações gradualmente adquirem diferenças genéticas. Consequentemente, todas as populações alcançam a fixação, algumas se tornam fixas para um alelo e outras se tornam fixas para um alelo alternativo.

Os três resultados da deriva genética (mudança da frequência alélica, perda da variação dentro das populações e divergência genética entre as populações) ocorrem simultaneamente e todos resultam do erro de amostragem. Os dois primeiros resultados ocorrem dentro das populações, enquanto o terceiro ocorre entre as populações. Resolva o Problema 35

Conceitos

A deriva genética é o resultado de uma população pequena contínua, o efeito do fundador (estabelecimento de uma população por alguns fundadores) e o efeito gargalo (redução da população). A deriva genética provoca uma mudança nas frequências alélicas dentro de uma população, uma perda da variação genética pela fixação dos alelos e a divergência genética entre as populações.

Seleção natural

Um processo final que produz mudanças nas frequências alélicas é a seleção natural, a reprodução diferencial dos genótipos (ver *Capítulo 24*). A seleção natural ocorre quando os indivíduos com traços adaptados produzem um número maior de descendentes do que os produzidos por outros na população. Se os traços adaptados têm uma base genética, eles são herdados pelos descendentes e aparecem com maior frequência na próxima geração. Um traço que fornece uma vantagem reprodutiva, portanto, aumenta com o passar do tempo, permitindo às populações que se tornem mais bem adequadas a seus ambientes – mais bem adaptadas. A seleção natural é única entre as forças evolutivas no sentido de que ela promove a adaptação.

Aptidão (fitness) e coeficiente de seleção. O efeito da seleção natural sobre o *pool* de genes de uma população depende dos valores de aptidão dos genótipos na população. A **aptidão (*fitness*)** é definida como o sucesso reprodutivo relativo de um genótipo. Aqui, o termo *relativo* é crítico: a aptidão é o sucesso reprodutivo de um genótipo comparado com os sucessos reprodutivos de outros genótipos na população. A aptidão (W) varia de 0 a 1. Suponha que o número médio de descendentes viáveis produzidos por três genótipos é

Genótipos	A^1A^1	A^1A^2	A^2A^2
Número médio de descendentes produzidos:	10	5	2

Para calcular a aptidão de cada genótipo, tomamos o número médio de descendentes produzidos por um genótipo e o dividimos pelo número médio de descendentes produzidos pelo genótipo mais fértil:

$$\begin{matrix} A^1A^1 & A^1A^2 & A^2A^2 \\ W_{11} = \frac{10}{10} = 1{,}0 & W_{12} = \frac{5}{10} = 0{,}5 & W_{22} = \frac{2}{10} = 0{,}2 \end{matrix} \qquad (25.17)$$

A aptidão do genótipo A^1A^1 é designada W_{11}, a aptidão de A^1A^2 é W_{12} e a aptidão de A^2A^2 é W_{22}. Uma variável relativa é o **coeficiente de seleção** (s), que é a intensidade relativa de seleção contra um genótipo. Geralmente falamos de seleção para determinado genótipo, mas tenha em mente que a seleção, quando é *a favor de* um genótipo, é automaticamente *contra* pelo menos um outro genótipo. O coeficiente de seleção é igual a $1 - W$, então os coeficientes de seleção para os três genótipos precedentes são:

	A^1A^1	A^1A^2	A^2A^2
Coeficiente de seleção ($1 - W$):	$s_{11} = 0$	$s_{12} = 0{,}5$	$s_{22} = 0{,}8$

Conceitos

A seleção natural é a reprodução diferencial de genótipos. É medida como aptidão, que é o sucesso reprodutivo de um genótipo comparado com outros genótipos em uma população.

Checagem dos conceitos 9

O número médio de descendentes viáveis produzidos por três genótipos é: *GG* = 6; *Gg* = 3, *gg* = 2. Qual é o valor da aptidão de *Gg*?

a. 3.
b. 0,5.
c. 0,3.
d. 0,27.

Modelo de seleção geral. Na seleção, a aptidão diferencial dos genótipos resulta em alteração das frequências dos genótipos com o tempo, que, por sua vez, provoca alterações nas frequências dos alelos que compõem os genótipos. Podemos prever o efeito da seleção natural sobre as frequências alélicas ao usar um modelo de seleção geral, que é descrito no **Quadro 25.4**. O uso desse modelo demanda conhecimento das frequências alélicas iniciais e dos valores de aptidão dos genótipos. Parte-se do pressuposto de que o acasalamento é aleatório e que a única força atuando em uma população é a seleção natural. O modelo de seleção geral pode ser usado para calcular as frequências alélicas após qualquer tipo de seleção. Também é possível desenvolver fórmulas para determinar a mudança na frequência alélica em que a seleção é contra traços recessivos, dominantes e codominantes, assim como os traços nos quais o heterozigoto tem aptidão maior (**Quadro 25.5**).

Conceitos

A alteração da frequência alélica em decorrência de seleção pode ser determinada para qualquer tipo de traço genético pelo uso do modelo de seleção geral.

Quadro 25.4 Método para determinar alterações da frequência alélica em decorrência de seleção.

	A^1A^1	A^1A^2	A^2A^2
Frequências genotípicas iniciais	p^2	$2pq$	q^2
Aptidão	W_{11}	W_{12}	W_{22}
Contribuição proporcional dos genótipos para a população	p^2W_{11}	$2pqW_{12}$	q^2W_{22}
Frequência genotípica relativa após a seleção	$\frac{p^2W_{11}}{\overline{w}}$	$\frac{2pqW_{12}}{\overline{w}}$	$\frac{q^2W_{22}}{\overline{w}}$

Nota: $w = p^2W_{11} + 2pqW_{12} + q^2W_{22}$.

Frequências alélicas após a seleção: $p' = f(A^1) = f(A^1A^1) + 1/2\ f(A^1A^2)$; $q' = 1 - p'$.

Quadro 25.5 Fórmulas para calcular a alteração das frequências alélicas com diferentes tipos de seleção.

	Valores de aptidão			
Tipo de seleção	A^1A^1	A^1A^2	A^2A^2	**Mudança em q**
Seleção contra um traço recessivo	1	1	$1 - s$	$\frac{-spq^2}{1 - sq^2}$
Seleção contra um traço dominante	1	$1 - s$	$1 - s$	$\frac{-spq^2}{1 - s + sq^2}$
Seleção contra um traço sem dominância	1	$1 - 1/2\ s$	$1 - s$	$\frac{-\frac{1}{2}spq}{1 - sq}$
Seleção contra ambos homozigotos (sobredominância)	$1 - s_{11}$	1	$1 - s_{22}$	$\frac{pq(s_{11}p - s_{22}q)}{1 - s_{11}p^2 - s_{22}q^2}$

Problema desenvolvido

Vamos aplicar o modelo de seleção geral no Quadro 25.4 para um exemplo real. O álcool é uma substância comum nas frutas podres, nas quais as larvas das mosca-da-fruta crescem e se desenvolvem; as larvas usam a enzima álcool desidrogenase (ADH) para desintoxicar os efeitos do álcool. Em algumas populações de mosca-da-fruta, dois alelos são encontrados no *locus* codificador de ADH: Adh^F, que codifica uma forma da enzima que migra rapidamente em um gel de eletroforese, e Adh^S, que codifica uma forma da enzima que migra lentamente em um gel de eletroforese. As moscas-da-fruta fêmeas com diferentes genótipos *Adh* produzem os seguintes números de descendentes quando existe álcool:

Genótipo	Número médio de descendentes
Adh^F/Adh^F	120
Adh^F/Adh^S	60
Adh^S/Adh^S	30

a. Calcule a aptidão relativa de fêmeas com estes genótipos.

b. Se uma população de mosca-da-fruta tem uma frequência inicial de Adh^F igual a 0,2, qual será a frequência na próxima geração quando houver álcool?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?

a. A aptidão das fêmeas de cada genótipo.

b. A frequência de Adh na próxima geração quando existe álcool.

Quais informações são fornecidas para solucionar o problema?

a. O número médio de descendentes produzidos por cada genótipo.

b. A população tem uma frequência alélica inicial de 0,2.

Etapas para a solução

a. Primeiro, é preciso calcular os valores de aptidão dos três genótipos. A aptidão (*fitness*) – o débito reprodutor relativo de um genótipo – é calculada ao se dividir o número médio de descendentes produzidos por esse genótipo pelo número médio de descendentes produzidos pelo genótipo mais fértil. Os valores de aptidão dos três genótipos *Adh* são:

Genótipo	Número médio de descendentes	Aptidão
Adh^F/Adh^F	120	$W_{FF} = \frac{120}{120} = 1$
Adh^F/Adh^S	60	$W_{FS} = \frac{60}{120} = 0,5$
Adh^S/Adh^S	30	$W_{SS} = \frac{30}{120} = 0,25$

b. Para calcular a frequência do alelo Adh^F após a seleção, podemos aplicar o método do quadro. Na primeira linha do quadro a seguir, registramos as frequências genotípicas iniciais antes de a seleção agir. Se o acasalamento é aleatório (uma suposição do modelo), os genótipos terão as frequências do equilíbrio de Hardy-Weinberg de p^2, $2pq$ e q^2. Na segunda linha do quadro a seguir, colocamos os valores de aptidão dos genótipos correspondentes. A proporção da população representada por cada genótipo após a seleção é obtida ao multiplicar a frequência genotípica inicial por sua aptidão (terceira linha do Quadro 25.4). Agora os genótipos não estão mais no equilíbrio de Hardy-Weinberg.

A aptidão média ($\bar{w}$) da população é a soma das contribuições proporcionais dos três genótipos: $\bar{w} = p^2W_{11} + 2pqW_{12} + q^2W_{22} = 0,04 + 0,16 + 0,16 = 0,36$. A aptidão média $\bar{w}$ é a aptidão média de todos os indivíduos na população e permite que sejam obtidos os genótipos após a seleção.

	Adh^FAdh^F	Adh^FAdh^S	Adh^SAdh^S
Frequências genotípicas iniciais:	$p^2 = (0,2)^2 = 0,04$	$2pq = 2(0,2)(0,8) = 0,32$	$q^2 = (0,8)^2 = 0,64$
Aptidão:	$W_{FF} = 1$	$W_{FS} = 0,5$	$W_{SS} = 0,25$
Contribuição proporcional de genótipos para população:	$p^2W_{FF} = 0,04(1) = 0,04$	$2pqW_{FS} = (0,32)(0,5) = 0,16$	$q^2W_{SS} = (0,64)(0,25) = 0,16$

A frequência de um genótipo após a seleção será igual a sua contribuição proporcional dividida pela aptidão média da população ($p^2W_{11}/\bar{w}$ para genótipo A^1A^1, $2pqW_{12}/\bar{w}$ para genótipo A^1A^2, e $q^2W_{22}/$ para genótipo A^2A^2), como demonstrado na quarta linha do Quadro 25.4. Agora podemos adicionar estes valores em nosso quadro a seguir:

	Adh^FAdh^F	Adh^FAdh^S	Adh^SAdh^S
Frequências genotípicas iniciais:	$p^2 = (0,2)^2 = 0,04$	$2pq = 2(0,2)(0,8) = 0,32$	$q^2 = (0,8)^2 = 0,64$
Aptidão:	$W_{FF} = 1$	$W_{FS} = 0,5$	$W_{SS} = 0,25$
Contribuição proporcional de genótipos para população:	$p^2W_{FF} = 0,04(1) = 0,04$	$2pqW_{FS} = (0,32)(0,5) = 0,16$	$q^2W_{SS} = (0,64)(0,25) = 0,16$

Frequência genotípica relativa após a seleção: $\frac{p^2 w_{FF}}{\bar{w}} = \frac{0,04}{0,36} = 0,11$ $\frac{2pqw_{FS}}{\bar{w}} = \frac{0,16}{0,36} = 0,44$ $\frac{q^2 w_{SS}}{\bar{w}} = \frac{0,16}{0,36} = 0,44$

Após as novas frequências genotípicas serem calculadas, a nova frequência alélica de Adh^F (p') pode ser determinada usando-se fórmula agora familiar da Equação 25.4:

$$p' = f(Adh^F) = f(Adh^F/Adh^F) + 1/2\, f(Adh^F/Adh^S) = 0,11 + 1/2(0,44) = 0,33$$

e a de q' pode ser obtida pela subtração:

$$q' = 1 - p' = 1 - 0,33 = 0,67$$

Prevemos que a frequência de Adh^F aumentará de 0,2 para 0,33,

Para praticar mais sobre o modelo de seleção, resolva o **Problema 37** no final deste capítulo.

Resultados da seleção. Os resultados da seleção dependem da aptidão relativa dos genótipos. Se tivermos três genótipos (A^1A^1, A^1A^2, A^2A^2) com valores de aptidão W_{11}, W_{12} e W_{22}, podemos identificar seus tipos diferentes de seleção natural (**Quadro 25.6**). Na seleção de tipo 1, um alelo dominante A^1 confere uma vantagem da aptidão; neste caso, a aptidão dos genótipos A^1A^1 e A^1A^2 é igual, e maior que a aptidão de A^2A^2 ($W_{11} = W_{12} > W_{22}$). Como o heterozigoto e o homozigoto A^1A^1 têm cópias do alelo A^1 e produzem mais descendentes do que o homozigoto A^2A^2, a frequência do alelo A^1 aumentará com o tempo e a frequência do alelo A^2 reduzirá. Esta forma de seleção, na qual um alelo ou traço é favorecido por outro, é chamada de **seleção direcional**.

A seleção de tipo 2 (ver Quadro 25.6) é a seleção direcional contra um alelo dominante A^1 ($W_{11} = W_{12} < W_{22}$). Neste caso, o alelo A^2 aumenta e o alelo A^1 diminui. As seleções tipo 3 e tipo 4 também são seleções direcionais, mas nestes casos, não existe dominância completa e o heterozigoto tem uma aptidão intermediária entre os dois homozigotos ($W_{11} > W_{12} > W_{22}$ para o tipo 3; $W_{11} < W_{12} < W_{22}$ para o tipo 4). Quando A^1A^1 tem a maior aptidão (tipo 3), o alelo A^1 aumenta e o alelo A^2 reduz-se com o passar do tempo. Quando A^2A^2 tem a maior aptidão (tipo 4), o alelo A^2 aumenta e o alelo A^1 reduz-se com o passar do tempo. Consequentemente, a seleção direcional leva à fixação do alelo favorecido e à eliminação de outro alelo, desde que nenhuma outra força evolutiva atue sobre a população.

Dois tipos de seleção (tipos 5 e 6) são situações especiais que levam ao equilíbrio, em que não existe mudança na frequência alélica. A seleção de tipo 5 é chamada de **sobredominância** ou vantagem do heterozigoto. Aqui, o heterozigoto tem maior aptidão do que os dois homozigotos ($W_{11} < W_{12} > W_{22}$). Com a sobredominância, ambos alelos são favorecidos no heterozigoto, e nenhum alelo é eliminado da população. Inicialmente, as frequências alélicas podem mudar porque um homozigoto tem maior aptidão do que outro; a direção da mudança dependerá dos valores de aptidão relativos dos dois homozigotos. As frequências alélicas mudam com a seleção sobredominante até um equilíbrio estável ser alcançado, no ponto em que não há mais mudança. A frequência alélica no equilíbrio ($\hat{q}$) depende dos valores de aptidão relativa (em geral expressados como coeficientes de seleção) dos dois homozigotos:

Quadro 25.6 Tipos de seleção natural.

Tipo	Relação de aptidão	Forma de seleção	Resultado
1	$W_{11} = W_{12} > W_{22}$	Seleção direcional contra o alelo recessivo A^2	A^1 aumenta, A^2 diminui
2	$W_{11} = W_{12} < W_{22}$	Seleção direcional contra o alelo dominante A^1	A^2 aumenta, A^1 diminui
3	$W_{11} > W_{12} > W_{22}$	Seleção direcional incompleta contra o alelo dominante A^2	A^1 aumenta, A^2 diminui
4	$W_{11} < W_{12} < W_{22}$	Seleção direcional incompleta contra o alelo dominante A^1	A^2 aumenta, A^1 diminui
5	$W_{11} < W_{12} > W_{22}$	Sobredominância	Equilíbrio estável, ambos alelos mantidos

6	$W_{11} > W_{12} < W_{22}$	Subdominância	Equilíbrio instável

Nota: W_{11}, W_{12} e W_{22} representam os valores de aptidão dos genótipos A^1A^1, A^1A^2 e A^2A^2, respectivamente.

$$\hat{q} = f(A^2) = \frac{s_{11}}{s_{11} + s_{22}} \tag{25.18}$$

Em que s_{11} representa o coeficiente de seleção do homozigoto A^1A^1 e s_{22} representa o coeficiente de seleção do homozigoto A^2A^2.

Um exemplo de sobredominância é o gene para anemia falciforme nos humanos, uma doença oriunda de uma mutação em um dos genes que codifica a hemoglobina. Pessoas homozigotas para a mutação falciforme produzem apenas a hemoglobina falciforme, apresentam grave anemia e, com frequência, lesão tecidual. As pessoas heterozigotas – com uma cópia normal e uma cópia mutante do gene – produzem a hemoglobina normal e a falciforme, mas suas hemácias têm hemoglobina normal suficiente para evitar a anemia falciforme. Entretanto, os heterozigotos são resistentes à malária e têm maior aptidão do que os homozigotos para hemoglobina normal e os homozigotos para anemia falciforme.

O último tipo de seleção (tipo 6) é **subdominância**, na qual o heterozigoto tem menor aptidão que ambos homozigotos ($W_{11} > W_{12} < W_{22}$). A subdominância leva a um equilíbrio instável, ou seja, as frequências alélicas não mudam enquanto estão em equilíbrio, mas se forem deslocadas do ponto de equilíbrio por alguma outra força evolutiva, elas se afastarão do equilíbrio até que um alelo acabe se tornando fixo. Resolva o Problema 38

Conceitos

A seleção natural modifica as frequências alélicas. O sentido e a magnitude da alteração dependem da intensidade da seleção, das relações de dominância dos alelos e das frequências alélicas. A seleção direcional favorece um alelo em vez de outro e acaba resultando em fixação do alelo favorecido. A sobredominância leva a um equilíbrio estável com manutenção dos alelos na população. A subdominância provoca um equilíbrio instável porque o heterozigoto tem menor aptidão do que os dois homozigotos.

Checagem dos conceitos 10

Qual a diferença da sobredominância para a seleção direcional?

Mudança na frequência alélica de um alelo recessivo devido à seleção natural. A taxa na qual a seleção muda as frequências alélicas depende da própria frequência alélica. Se um alelo (A^2) é letal e recessivo, $W_{11} = W_{12} = 1$, enquanto $W_{22} = 0$. A frequência do alelo A^2 se reduzirá com o tempo (porque o homozigoto A^2A^2 não produz descendentes) e a taxa de redução será proporcional à frequência do alelo recessivo. Quando a frequência do alelo é alta, a mudança em cada geração é relativamente grande, mas, à medida que a frequência do alelo cai, maior proporção de alelos está nos genótipos heterozigotos, nos quais eles são imunes à ação da seleção natural (os heterozigotos têm o mesmo fenótipo que o homozigoto favorecido). Assim, a seleção contra um alelo recessivo raro é muito ineficiente e sua remoção da população é lenta.

A relação entre a frequência de um alelo recessivo e sua taxa de mudança pela seleção natural tem uma importante implicação. Algumas pessoas acreditam que o tratamento médico de pacientes com doenças recessivas raras acabará levando a aumento do gene da doença, levando por fim à degeneração do *pool* de genes humanos. Essa crença equivocada foi a base de leis eugênicas implantadas no início do século 20 proibindo o casamento de pessoas com determinadas condições genéticas e permitindo a esterilização involuntária de outras. Entretanto, a maioria das cópias de alelos recessivos raros ocorre nos heterozigotos, e a seleção contra os homozigotos terá pouco efeito sobre a frequência de um alelo recessivo. Assim, se os homozigotos para um traço recessivo se reproduzirem ou não tem pouco efeito sobre a frequência do distúrbio.

Mutação e seleção natural. A mutação recorrente e a seleção natural atuam como forças opostas nos alelos prejudiciais; a mutação aumenta sua frequência e a seleção natural reduz sua frequência. Consequentemente, estas duas forças alcançam um equilíbrio, no qual o número de alelos adicionados pela mutação é balanceado pelo número de alelos removidos pela seleção.

A frequência de um alelo recessivo prejudicial no equilíbrio ($\hat{q}$) é igual à raiz quadrada da taxa de mutação dividida pelo coeficiente de seleção:

(25.19)

$$\hat{q} = \sqrt{\frac{\mu}{s}}$$

Para que a seleção atue em um alelo dominante prejudicial, a frequência do alelo dominante no equilíbrio pode ser apresentada como:

$$\hat{q} = \frac{\mu}{s} \tag{25.20}$$

A acondroplasia é um tipo comum de nanismo humano causado por um gene dominante. As pessoas com esta condição são férteis, embora tenham apenas 74% do número de filhos gerados por pessoas sem acondroplasia. A aptidão das pessoas com acondroplasia é, portanto, em média 0,74, e o coeficiente de seleção (s) é 1 – *W* ou 0,26. Se pressupormos que a taxa de mutação para acondroplasia é cerca de 3×10^{-5} (uma típica taxa de mutação nos humanos), então podemos prever que a frequência de equilíbrio para o alelo da acondroplasia será

$$\hat{q} = \frac{0{,}00003}{0{,}26} = 0{,}0001153$$

Esta frequência está perto da frequência real da condição. **Resolva o Problema 41**

Conceitos

A mutação e a seleção natural atuam como forças opostas nos alelos prejudiciais: a mutação aumenta sua frequência e a seleção natural reduz sua frequência, produzindo um equilíbrio.

Conceitos conectantes

Efeitos gerais das forças que mudam as frequências alélicas

Agora você sabe que quatro processos modificam as frequências alélicas de uma população: mutação, migração, deriva genética e seleção natural. Seus efeitos de longo e curto prazos sobre as frequências alélicas estão resumidos no **Quadro 25.7**. Em alguns casos, estas mudanças continuam até um alelo ser eliminado e o outro se tornar fixo na população. A deriva genética e a seleção direcional resultarão na fixação, desde que estas forças sejam as únicas que atuem em uma população. Com as outras forças evolutivas, as frequências alélicas mudam até alcançar um ponto de equilíbrio e não há mudança adicional na frequência alélica. Mutação, migração e algumas formas de seleção natural podem levar aos equilíbrios estáveis (ver Quadro 25.7).

Diferentes forças evolutivas afetam a variação genética dentro das populações e a divergência genética entre as populações. As forças evolutivas que mantêm ou reduzem a variação genética intrapopulacional estão listadas no quadrante superior à esquerda da **Figura 25.12**. Estas forças incluem alguns tipos de seleção natural, como a sobredominância, na qual ambos alelos são favorecidos. A mutação e a migração também aumentam a variação genética dentro das populações porque elas introduzem novos alelos na população. As forças evolutivas que reduzem a variação genética dentro das populações estão listadas no quadrante inferior à esquerda da Figura 25.12. Estas forças incluem a deriva genética, que reduz a variação pela fixação dos alelos e algumas formas de seleção natural como a seleção direcional.

As várias forças evolutivas também afetam a quantidade da divergência genética entre as populações. A seleção natural aumenta a divergência entre as populações se diferentes alelos forem favorecidos em diferentes populações, mas ela também reduz a divergência entre as populações ao favorecer o mesmo alelo nas diferentes populações. A mutação quase sempre aumenta a divergência entre as populações porque diferentes mutações surgem em cada população. A deriva genética também aumenta a divergência entre as populações porque mudanças nas frequências alélicas devido à deriva são aleatórias e é provável que mudem nas diferentes direções em populações separadas. A migração, por outro lado, reduz a divergência entre as populações porque ela torna a composição genética semelhante entre as populações.

A migração e a deriva genética atuam em direções opostas: a migração aumenta a variação genética dentro das populações e reduz a divergência entre as populações, enquanto a deriva genética aumenta a

variação genética dentro das populações e aumenta a divergência entre as populações. A mutação aumenta a variação dentro de populações e divergência entre populações. A seleção natural pode aumentar ou reduzir a variação entre as populações e pode aumentar ou reduzir a divergência entre as populações.

Um ponto importante a ter em mente é que as populações reais são simultaneamente afetadas por muitas forças evolutivas. Examinamos os efeitos da mutação, migração, deriva genética e seleção natural no isolamento de modo que a influência de cada processo estaria clara. Entretanto, no mundo real, as populações são afetadas por várias forças evolutivas ao mesmo tempo e a evolução resulta da ação complexa de numerosos processos.

Quadro 25.7 Efeitos de diferentes forças evolutivas sobre as frequências alélicas dentro das populações.

Força	Efeito de curto prazo	Efeito de longo prazo
Mutação	Mudança na frequência alélica	Equilíbrio alcançado entre as mutações direta e reversa
Migração	Mudança na frequência alélica	Equilíbrio alcançado quando as frequências alélicas da população fonte e receptora são iguais
Deriva genética	Mudança na frequência alélica	Fixação de um alelo
Seleção natural	Mudança na frequência alélica	Seleção direcional: fixação de um alelo Seleção sobredominante: equilíbrio alcançado

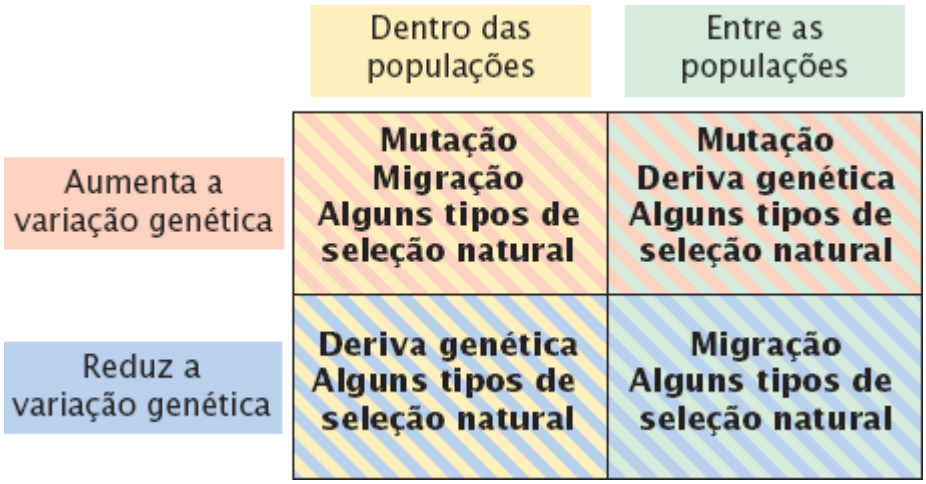


Figura 25.12 Mutação, migração, deriva genética e seleção natural têm efeitos diferentes sobre a variação genética dentro das populações e sobre a divergência genética entre as populações.

Resumo dos conceitos

- A genética de populações examina a composição genética de uma população e como esta composição muda com o tempo
- Uma população mendeliana é um grupo de indivíduos intercruzados, sexualmente reprodutivos, cujo conjunto de genes constitui o *pool* de genes da população. A evolução ocorre por mudanças neste *pool* de genes
- A composição genética de uma população pode ser descrita por suas frequências genotípicas e alélicas
- A lei de Hardy-Weinberg descreve os efeitos da reprodução e das leis de Mendel sobre as frequências alélicas e genotípicas de uma população. Ela considera que uma população é grande, com acasalamento aleatório e livre dos efeitos da mutação, migração e seleção natural. Quando estas condições são satisfeitas, as frequências alélicas não mudam e as frequências genotípicas estabilizam-se após uma geração na população do equilíbrio de Hardy-Weinberg nas proporções p^2, $2pq$ e q^2, em que p e q têm frequências iguais dos alelos
- O acasalamento não aleatório afeta as frequências dos genótipos, mas não as frequências dos alelos
- A endogamia, um tipo de acasalamento preferencial, aumenta a frequência dos homozigotos enquanto reduz a frequência de heterozigotos. A endogamia é prejudicial porque aumenta o surgimento de traços recessivos letais e prejudiciais

- Mutação, migração, deriva genética e seleção natural podem mudar as frequências alélicas
- A mutação recorrente leva a um equilíbrio, com as frequências alélicas sendo determinadas pelas taxas relativas de mutação direta e reversa. A mudança pela mutação em uma única geração em geral é muito pequena porque as taxas de mutação são baixas
- A migração, o movimento dos genes entre populações, aumenta a quantidade de variação genética dentro das populações e reduz o número de diferenças entre as populações
- A deriva genética é a mudança na frequência alélica causada por fatores do acaso. A deriva genética aumenta quando uma população composta por um pequeno número de indivíduos é estabelecida por poucos fundadores, ou sofre uma grande redução de tamanho. A deriva genética muda as frequências alélicas, reduz a variação genética dentro das populações e provoca divergência genética entre as populações
- A seleção natural é a reprodução diferencial dos genótipos; é medida pelos relativos sucessos de reprodução (aptidão) dos genótipos. Os efeitos da seleção natural sobre a frequência alélica podem ser determinados ao aplicar o modelo de seleção geral. A seleção direcional leva à fixação de um alelo. A taxa de mudança na frequência alélica devido à seleção depende da intensidade da seleção, das relações de dominância e das frequências iniciais dos alelos
- A mutação e a seleção natural podem produzir um equilíbrio, no qual o número de alelos introduzidos pela mutação é balanceado pela eliminação dos alelos pela seleção natural.

Termos importantes

Acasalamento preferencial negativo
Acasalamento preferencial positivo
Aptidão (*fitness*)
Coeficiente de endogamia
Coeficiente de seleção
Depressão da endogamia
Deriva genética
Efeito do fundador
Endogamia
Equilíbrio
Equilíbrio de Hardy-Weinberg
Erro de amostragem
Exocruzamento (*outcrossing*)
Fixação
Frequência alélica
Frequência genotípica
Gargalo genético
Lei de Hardy-Weinberg
Migração (fluxo gênico)
Pool de genes
População mendeliana
Resgate genético
Seleção direcional
Sobredominância
Subdominância
Tamanho efetivo da população

Respostas da Checagem dos conceitos

1. Existem menos alelos do que genótipos, então o *pool* de genes de uma população pode ser descrito em poucos parâmetros quando as frequências alélicas são usadas. Além disso, os genótipos são junções temporárias de alelos que degradam em cada geração; os alelos são transmitidos de geração em geração nos organismos de reprodução sexuada.
2. a.
3. c.
4. 0,10.
5. b.

6. c.
7. c.
8. d.
9. b.
10. Na sobredominância, o heterozigoto tem a maior aptidão e ambos alelos são mantidos. Na seleção direcional, a seleção faz com que um alelo ou traço tenha maior frequência.

Problema desenvolvido

Problema

Um alelo recessivo para cabelo ruivo (*r*) tem uma frequência de 0,2 na população I e uma frequência de 0,01 na população II. A fome na população I faz com que várias pessoas na população I migrem para população II, onde se reproduzem aleatoriamente com os membros da população II. Os geneticistas estimam que, após a migração, 15% das pessoas na população II serão pessoas que migraram da população I. Qual será a frequência de cabelo ruivo na população II após a migração?

Estratégias para a solução

De quais informações você precisa para solucionar o problema?
A frequência de cabelo ruivo na população II após a migração.

Quais informações são fornecidas para solucionar o problema?

- O alelo para cabelo ruivo (*r*) é recessivo
- A frequência inicial do *r* na população I é 0,2
- A frequência inicial do *r* na população II é 0,01
- Após a migração, 15% das pessoas na população II são pessoas que migraram da população I.

Para a solução deste problema, revise:
Migração, na Seção 25.4.

Etapas para a solução

A partir da Equação 25.14, a frequência alélica em uma população após a migração (q'_{II}) é

$$q'_{II} = q_I\,(m) + q_{II}\,(1 - m)$$

Em que q_I e q_{II} são as frequências alélicas na população I (migrantes) e na população II (residentes), respectivamente e m é a proporção da população II composta de migrantes. Neste problema, a frequência do cabelo ruivo é 0,2 na população I e 0,01 na população II. Como 15% da população II é composta por migrantes, $m = 0{,}15$. Substituindo estes valores na Equação 25.14, temos

$$q'_{II} = 0{,}2(0{,}15) + (0{,}01)(1 - 0{,}15) = 0{,}03 + 0{,}0085 = 0{,}0385$$

que é a frequência esperada do alelo para cabelo ruivo na população II após a migração. O cabelo ruivo é um traço recessivo; se o acasalamento para o cabelo ruivo for aleatório, a frequência desta cor de cabelo na população II após a migração será

$$f\,(rr) = q^2 = (0{,}0385)^2 = 0{,}0015$$

Lembrete: Com o acasalamento aleatório, as frequências esperadas dos genótipos são p^2 *(RR)*, $2pq$ *(Rr)* e q^2 (*rr*).

Questões de compreensão

Seção 25.1

1. O que é uma população mendeliana? Como o *pool* de genes de uma população mendeliana é descrito?

Seção 25.2

2. Quais são as previsões fornecidas pela lei de Hardy-Weinberg?
3. Quais suposições devem ser satisfeitas para uma população estar no equilíbrio de Hardy-Weinberg?
4. O que é acasalamento aleatório?
5. Indique as frequências genotípicas de Hardy-Weinberg esperadas para (a) um *locus* autossômico com três alelos e (b) um *locus* ligado ao X com dois alelos.

Seção 25.3

6. Defina a endogamia e descreva resumidamente seus efeitos em uma população.

Seção 25.4

7. O que determina as frequências alélicas no equilíbrio mutacional?
8. Quais fatores afetam a magnitude de mudança nas frequências alélicas causada pela migração?
9. Defina deriva genética e indique três formas nas quais ela pode surgir. Qual efeito a deriva genética exerce sobre uma população?
10. Qual seu efeito no tamanho da população? Como ela afeta a quantidade de deriva genética?
11. Defina seleção natural e aptidão.
12. Descreva resumidamente as diferenças entre seleção direcional, sobredominância e subdominância. Descreva o efeito de cada tipo de seleção sobre as frequências alélicas de uma população.
13. Quais fatores afetam a taxa de mudança na frequência alélica causada pela seleção natural?
14. Compare e contraste os efeitos da mutação, migração, deriva genética sobre a variação genética dentro das populações e sobre a divergência genética entre as populações.

Questões e problemas aplicados

Seção 25.1

15. Como você responderia a uma pessoa que diz que os modelos são inúteis no estudo da genética de populações porque eles representam grandes simplificações do mundo real?

*16. Arganazes (*Microtus ochrogaster*) foram presos nos campos no sul de Indiana e genotipados para um *locus* da transferrina. Foram registrados os números de genótipos a seguir, em que T^E e T^F representam diferentes alelos.

T^ET^E	T^ET^F	T^FT^F
407	170	17

Calcule as frequências genotípica e alélica no *locus* da transferrina para esta população.

17. Jean Manning, Charles Kerfoot e Edward Berger estudaram as frequências genotípicas do *locus* da fosfoglicose isomerase (GPI) no crustáceo *Bosmina longirostris* (um pequeno crustáceo conhecido como pulga-d'água). Em uma localidade, eles coletaram 176 animais da Union Bay em Seattle, Washington e determinaram os seus genótipos de GPI usando eletroforese (J. Manning, W. C. Kerfoot, and E. M. Berger. 1978. *Evolution* 32:365-374).

Genótipo	Número
S^1S^1	4
S^1S^2	38
S^2S^2	134

Determine as frequências genotípica e alélica para esta população.

*18. A cor de pelo laranja dos gatos é causada por um alelo ligado ao X (X^O) que é codominante com o alelo para preto (X^+). Foram determinados os genótipos do *locus* laranja dos gatos em Minneapolis e St. Paul, Minnesota e foram obtidos os dados a seguir:

Fêmeas X^OX^O	11
Fêmeas X^OX^+	70

Fêmeas $X^+ X^+$	94
Machos X^O Y	36
Machos X^+ Y	112

Calcule as frequências dos alelos X^O e X^+ para esta população.

Seção 25.2

19. Use o gráfico apresentado na Figura 25.3 para determinar qual genótipo é mais frequente quando a frequência do alelo *A* é:
    a. 0,2. b. 0,5. c. 0,8.
20. Um total de 6.129 caucasianos norte-americanos tiveram seu tipo sanguíneo determinado para o *locus* MN, que é determinado por dois alelos codominantes L^M e L^N. Foram obtidos os seguintes dados:

Tipo sanguíneo	Número
M	1.787
MN	3.039
N	1.303

Faça um teste qui-quadrado para determinar se esta população está no equilíbrio de Hardy-Weinberg no *locus* MN.

21. Considere que os fenótipos de joaninhas apresentados na Figura 25.1 sejam codificados pelos seguintes fenótipos:

Fenótipo	Genótipo
Todas pretas	*BB*
Algumas manchas pretas	*Bb*
Sem manchas pretas	*bb*

    a. Para as joaninhas apresentadas na figura, calcule as frequências dos genótipos e as frequências dos alelos.
    b. Use o teste qui-quadrado para determinar se as joaninhas apresentadas aqui estão no equilíbrio de Hardy-Weinberg.

ANÁLISE DE DADOS

*22. A maioria dos ursos-negros (*Ursos americanus*) é preta ou marrom. Entretanto, aparecem ursos brancos desta espécie em algumas populações ao longo da costa da British Columbia. Kermit Ritland e colaboradores determinaram que a cor branca nestes ursos é o resultado de uma mutação recessiva (G) causada por uma substituição de nucleotídio único no qual a guanina é substituída pela adenina no *locus* do receptor da melanocortina-1 (*mrc1*), o mesmo *locus* responsável pelo cabelo ruivo nos humanos (K. Ritland, C. Newton, and H. D. Marshall. 2001. *Current Biology* 11:1468-1472). O alelo do tipo selvagem neste *locus* (*A*) codifica a cor preta ou marrom. Ritland e colaboradores coletaram amostras dos ursos nas três ilhas e determinaram seus genótipos no *locus mcr1*.

Genótipo	Número
AA	42
AG	24
GG	21

    a. Quais são as frequências dos alelos *A* e *G* nestes ursos?
    b. Indique as frequências genotípicas esperadas se a população está no equilíbrio de Hardy-Weinberg.
    c. Use um teste qui-quadrado para comparar os números dos genótipos observados com o número esperado sob o equilíbrio de Hardy-Weinberg. Esta população está no equilíbrio de Hardy-Weinberg? Explique sua justificativa.

23. Os genótipos das rãs-leopardo de uma população na região central do Kansas foram determinados para um *locus* (*m*) que codifica a enzima malato desidrogenase. Foram observados os seguintes números de genótipos:

Genótipo	Número
M^1M^1	20
M^1M^2	45
M^2M^2	42
M^1M^3	4
M^2M^3	8
M^3M^3	6

a. Calcule as frequências genotípica e alélica para esta população.
b. Quais seriam os números esperados de genótipos se a população estiver no equilíbrio de Hardy-Weinberg?

24. A cor uniforme (*D*) nos gatos domésticos é dominante sobre a cor não uniforme (*d*). Dos 325 gatos observados, 194 tinham cor uniforme e 131 tinham cor não uniforme.
    a. Se estes gatos estão no equilíbrio de Hardy-Weinberg para o *locus* não uniforme, qual é a frequência do alelo não uniforme?
    b. Quais dos 194 gatos com cor uniforme devem ser heterozigotos?

*25. A doença de Tay-Sachs é um distúrbio autossômico recessivo. Entre os judeus asquenazes, a frequência da doença de Tay-Sachs é 1 em 3.600. Se a população de asquenazes acasala aleatoriamente para o gene Tay-Sachs, qual a proporção da população é composta dos carreadores heterozigotos do alelo de Tay-Sachs?

26. Na planta *Lotus corniculatus*, o glicosídio cianogênico protege a planta contra as pragas de insetos e até o pasto pelo gado. Este glicosídio é produzido por um único alelo dominante. Uma população de *L. corniculatus* é composta por 77 plantas com o glicosídio cianogênico e 56 plantas sem o composto. Qual é a frequência do alelo dominante responsável pela presença do glicosídio cianogênico nesta população?

27. O daltonismo nos seres humanos é um traço recessivo ligado ao X. Aproximadamente 10% dos homens em uma população são daltônicos.
    a. Se o acasalamento é aleatório para o *locus* do daltonismo, qual é a frequência do alelo do daltonismo nesta população?
    b. Qual proporção esperada das mulheres nesta população é daltônica?
    c. Qual proporção esperada das mulheres nesta população é carreadora heterozigota do alelo do daltonismo?

Seção 25.3

*28. O tipo sanguíneo MN humano é determinado por dois alelos codominantes L^M e L^N. A frequência de L^M nos esquimós de uma pequena ilha no Ártico é 0,80.
    a. Se ocorre acasalamento aleatório nesta população, quais são as frequências esperadas dos tipos sanguíneos M, MN e N na ilha?
    b. Se o coeficiente de endogamia nesta população é 0,50, quais são as frequências esperadas dos tipos sanguíneos M, MN e N na ilha?

29. Demonstre matematicamente que o acasalamento entre irmãos consaguíneos (F = 1/4) reduz a heterozigosidade em 1/4 a cada geração.

Seção 25.4

*30. A taxa de mutação direta para o malhado nos porquinhos-da-índia é 8×10^{-5}; a taxa de mutação reversa é 2×10^{-6}. Se considerarmos que não existem outras forças evolutivas, qual é a frequência esperada do alelo para malhado em uma população que está em equilíbrio mutacional?

ANÁLISE DE DADOS

31. Por 3 anos, Gunther Schlager e Margaret Dickie estimaram as taxas de mutação direta e reversa para cinco *loci* nos camundongos que codificam vários aspectos da cor do pelo ao examinar mais de 5 milhões de camundongos para mutações espontâneas (G. Schlager and M. M. Dickie 1966. *Science* 151:205-206). O número de mutações detectadas no *locus* não uniforme era:

	Número de gametas examinados	Número de mutações detectadas
Mutações diretas	260.675	5
Mutações reversas	583.360	2

Calcule as taxas de mutação direta e reversa neste *locus*. Se estas taxas de mutação são representativas das taxas nas populações naturais de camundongos, qual seria a frequência do equilíbrio esperada das mutações diluídas?

32. Na Figura 25.9, cada ponto azul representa uma cópia do alelo *A* e cada ponto vermelho representa cada cópia do alelo *a*. Calcule as frequências do alelo *A* na população II antes e após a migração. Explique por que a frequência de *A* na população II mudou após a migração.

33. Nas baratinhas, a asa curva (*cv*) é recessiva para asa normal (cv^+). Bill, que está criando baratinhas no seu quarto, descobre que a frequência do gene para as asas curvas na sua população de baratas é 0,6. No apartamento do seu amigo Joe, a

frequência do gene para asas curvas é 0,2. Um dia, Joe visita o quarto do Bill e várias baratinhas pulam do cabelo de Joe e se juntam à população do quarto do Bill. Agora Bill estima que 10% das baratinhas em seu quarto são as que pularam do cabelo de Joe. Qual é a nova frequência de asas curvas entre as baratinhas do quarto de Bill?

*34. Uma população de cobras-d'água é encontrada em uma ilha no Lago Erie. Algumas são listradas e outras não; a listra é causada por um alelo autossômico recessivo para um alelo para sem listras. A frequência de cobras listradas na ilha é 0,4, enquanto a frequência de cobras listradas no continente é 0,81. No verão, várias cobras migram do continente para a ilha. Após esta migração 20% da população da ilha são compostos por cobras oriundas do continente.

a. Se consideramos que a população do continente e da ilha estão no equilíbrio de Hardy-Weinberg para os alelos que afetam a listra, qual é a frequência do alelo para listras na ilha e no continente antes da migração?
b. Após ocorrer a migração, qual é a frequência do alelo de listras na ilha?

*35. Os pikas são pequenos mamíferos que vivem em altitudes elevadas nas encostas das montanhas. A maioria das populações localizadas nos topos das montanhas em Colorado e Montana na América do Norte está isolada uma da outra: os pikas não ocupam *habitats* de baixas altitudes que separam os topos das montanhas e não se arriscam para longe das altitudes elevadas. Assim, há pouco fluxo de genes entre as populações. Além disso, cada população é pequena e foi fundada por alguns poucos pikas.

Um grupo de geneticistas de populações se propôs a estudar a quantidade de variação genética em uma série de populações de pika e comparar as frequências alélicas em diferentes populações. Com base na biologia e na distribuição de pikas, preveja o que os geneticistas populacionais vão descobrir sobre a variação genética dentro e entre as populações.

36. Qual proporção das populações apresentadas na Figura 25.11 alcançou a fixação nas gerações 10, 25 e 30? Como a proporção de populações que alcançaram a fixação por deriva genética muda com o tempo?

*37. Em uma grande população de acasalamento aleatório, a frequência do alelo (*s*) para hemoglobina falciforme é 0,028. Os resultados dos estudos demonstram que as pessoas com os seguintes genótipos no *locus* da cadeia beta produzem os números médios de descendentes apresentados:

Genótipo	Número médio de descendentes produzidos
SS	5
Ss	6
ss	0

a. Qual será a frequência do alelo (*s*) de hemoglobina falciforme na próxima geração?
b. Qual será a frequência do alelo (*s*) de hemoglobina falciforme no equilíbrio?

*38. Duas inversões cromossômicas são comumente encontradas nas populações da *Drosophila pseudoobscura*: Padrão (*ST*) e Ponta de flecha (*AR*). Quando tratada com o inseticida DDT, os genótipos para estas inversões exibem sobredominância, com os seguintes valores de aptidão:

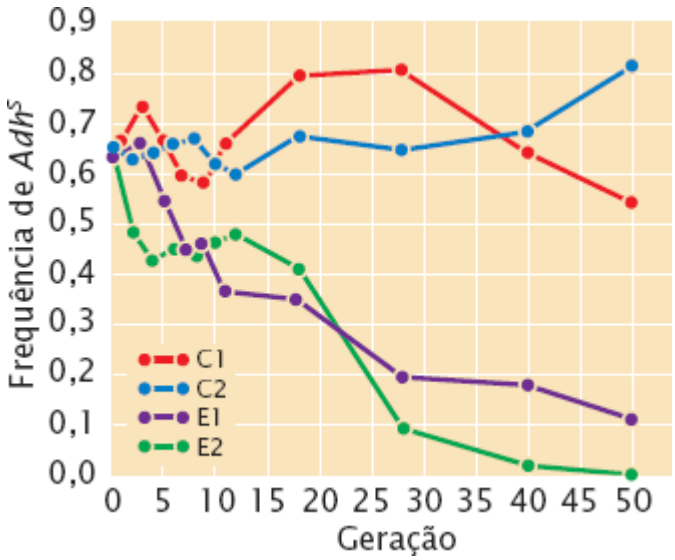
Genótipo	Aptidão
ST/ST	0,47
ST/AR	1
AR/AR	0,62

Quais serão as frequências de *ST* e *AR* após o equilíbrio ser alcançado?

39. Em uma grande população de acasalamento aleatório, a frequência de um alelo autossômico recessivo letal é 0,20. Qual será a frequência deste alelo na próxima geração se a letalidade ocorrer antes da reprodução?

ANÁLISE DE DADOS

40. A mosca-da-fruta *Drosophila melanogaster* se alimenta de fruta podre, que pode fermentar e ter altos níveis de álcool. Douglas Cavener e Michael Clegg estudaram as frequências alélicas no *locus* para álcool desidrogenase (*Adh*) em populações experimentais de *D. melanogaster* (D. R. Cavener e M. T. Clegg. 1981. *Evolution* 35:1-10). As populações experimentais foram estabelecidas a partir de moscas selvagens capturadas e foram criadas em gaiolas no laboratório. Duas populações (C1 e C2) foram criadas em uma dieta padrão de ágar melaço-farinha de milho. Duas populações de etanol (E1 e E2) foram criadas em uma dieta de melaço-farinha de milho à qual foram adicionados 10% de etanol. As quatro populações foram periodicamente amostradas para determinar as frequências alélicas dos dois alelos no *locus* da álcool desidrogenase, Adh^S e Adh^F. As frequências desses alelos nas populações experimentais estão apresentadas no gráfico.



a. Com base nestes dados, que conclusão você pode tirar sobre as forças evolutivas que afetam os alelos *Adh* nestas populações?

b. Cavener e Clegg mediram a viabilidade de diferentes genótipos *Adh* no ambiente com álcool e obtiveram os seguintes valores:

Genótipo	Viabilidade relativa
Adh^F/Adh^F	0,932
Adh^F/Adh^S	1,288
Adh^S/Adh^S	0,596

Usando essas viabilidades relativas, calcule a aptidão (*fitness*) relativa dos três genótipos. Se uma população tiver uma frequência inicial de $p = f(Adh^F) = 0,5$, qual será a frequência esperada de Adh^F na geração seguinte com bases nestes valores de aptidão?

*41. Um tipo de glaucoma congênito é causado por um alelo autossômico recessivo. Considere que a taxa de mutação é 10^{-5} e que pessoas com esta condição têm, em média, apenas cerca de 80% dos descendentes gerados por pessoas sem glaucoma.

a. No equilíbrio entre mutação e seleção, qual será a frequência do gene para o glaucoma congênito?

b. Qual será a frequência da doença em uma população de acasalamento aleatório que esteja em equilíbrio?

42. Examine a Figura 25.12. Quais forças evolutivas:

a. Provocam aumento da variação genética intrapopulacional e interpopulacional?

b. Provocam redução da variação genética intrapopulacional e interpopulacional?

c. Provocam aumento da variação genética intrapopulacional, mas provocam redução da variação genética interpopulacional?

Questões desafiadoras

Seção 25.2

43. A frequência do alelo *A* em uma população é 0,8 e a frequência do alelo *a* é 0,2. Se a população acasalar aleatoriamente para este *locus*, indique todos os possíveis acasalamentos entre os genótipos neste *locus* e a proporção esperada de cada tipo.

Seção 25.4

44. A salamandra *Eurycea sosorum* é uma espécie ameaçada de extinção encontrada apenas em três nascentes adjacentes na cidade de Austin, Texas. Existe a preocupação crescente de que o despejo de produtos químicos em uma autoestrada próxima poderia poluir a nascente e dizimar a espécie. Para conseguir uma fonte de salamandras para repovoar a nascente no caso de tal catástrofe, foi feita uma proposta para estabelecer uma população com criação em cativeiro da salamandra em um zoológico local. Pediram para você criar um plano para estabelecer esta população com criação em cativeiro, com o objetivo de manter o máximo de variação genética da espécie. Quais fatores poderiam levar a perda de variação genética na criação da população em cativeiro? Como a perda de tal variação poderia ser evitada? Com a suposição de que apenas um número limitado de salamandras pode ser mantido em cativeiro, quais procedimentos devem ser

instituídos para garantir a manutenção a longo prazo do máximo de variação?

26

Genética Evolutiva

Os genes do paladar nos macacos que cospem

Quase todo estudante de biologia conhece o teste do paladar. O professor passa pequenos pedaços de papel impregnados com um composto chamado feniltiocarbamida (PTC) e os estudantes, seguindo as instruções do professor, colocam o papel na sua boca. A reação é quase sempre a mesma: vários estudantes imediatamente cospem o papel, por causa do gosto amargo da PTC. Alguns alunos, entretanto, não sentem o gosto da PTC e continuam a sugar o papel, perguntando por que os outros alunos estão cuspindo. A variação entre os indivíduos em um traço como a capacidade de sentir o gosto da PTC é chamada de polimorfismo.

A capacidade de sentir o gosto da PTC é herdada como um traço autossômico dominante nos humanos. As frequências de alelos de provador e não provador foram estimadas em centenas de populações humanas pelo mundo. Quase todas as populações têm provadores e não provadores; a frequência dos dois alelos varia muito.

A PTC não é encontrada na natureza, mas a capacidade de sentir seu gosto está fortemente correlacionada com a capacidade de sentir gosto de outros compostos naturais de gosto amargo, alguns dos quais são tóxicos. A capacidade de sentir o gosto de PTC está ligada a preferências alimentares e pode estar associada com a suscetibilidade a algumas doenças, como deficiência da tireoide. Estas observações sugerem que a seleção natural participa na evolução do traço do provador (paladar). Conseguimos compreender parte da evolução do provador (paladar) quando o famoso geneticista de populações Ronald A. Fisher e colaboradores visitaram o zoológico em 1939.

Fisher queria saber se outros primatas também têm a capacidade de sentir o gosto da PTC. Para responder a esta pergunta, ele preparou algumas bebidas com diferentes concentrações de PTC e foi para o zoológico com seus amigos, os biólogos Edmund (E. B.) Ford e Julian Huxley. No zoológico, as oito bebidas com PTC foram oferecidas a oito chimpanzés e um orangotango. Fisher e seus amigos estavam preocupados que eles não fossem capazes de dizer se os macacos poderiam sentir o gosto da PTC. Esta preocupação desapareceu quando o primeiro que bebeu a bebida imediatamente cuspiu em Fisher. Dos oito chimpanzés testados, seis eram provadores e dois não eram.

A observação de que os chimpanzés e humanos têm o polimorfismo para o paladar de PTC levou Fisher e seus amigos a supor que o polimorfismo surgiu em um ancestral comum de humanos e chimpanzés, que o transmitiu para ambas as espécies. Entretanto, eles não tinham como testar sua hipótese. Sessenta e cinco anos depois, os geneticistas, munidos com as mais modernas técnicas de genética molecular, conseguiram determinar a origem do polimorfismo do paladar de PTC e testaram a hipótese de Fisher e seus amigos.

Estudos moleculares revelaram que nossa capacidade para sentir o paladar de PTC é controlada pelos alelos no *locus* *TAS2R38*, um gene de 1.000 pb encontrado no cromossomo 7. Este *locus* codifica receptores para compostos amargos e é expresso nas células das nossas papilas gustativas. Um alelo comum codifica um receptor que permite a capacidade de sentir o paladar de PTC, um alelo alternativo codifica um receptor que não responde a PTC.

Pesquisa recente demonstrou que a sensibilidade ao gosto de PTC nos chimpanzés também é controlada pelos alelos no *locus* *TAS2R38*. Entretanto, para surpresa dos investigadores os alelos de paladar nos humanos e chimpanzés não são iguais no nível molecular. Nos humanos com alelos de provador e não provador, as diferenças de nucleotídios em três posições afetam quais aminoácidos estão presentes em uma proteína receptora de paladar. Os chimpanzés não têm nenhuma destas diferenças nos nucleotídios. Em vez disso, uma mutação no códon de iniciação produz o alelo não provador. Esta substituição elimina o códon de iniciação normal, e o ribossomo inicia a tradução em um códon de iniciação alternativo *downstream*, resultando na produção de um receptor encurtado de proteína que não responde a PTC.

O que estas descobertas significam é que Fisher e seus amigos estavam corretos sobre humanos e chimpanzés terem o polimorfismo do paladar de PTC, mas estavam errados na sua hipótese sobre sua origem: os humanos e os chimpanzés desenvolveram de forma independente o polimorfismo para o paladar de PTC.

ste capítulo discute a base genética da evolução. Como ilustrado pelo polimorfismo do provador de PTC, a genética

Evolutiva tem uma longa história, mas recentemente foi transformada pela aplicação de poderosas técnicas de genética molecular. No Capítulo 25, analisamos as quatro forças evolutivas que modificam as frequências alélicas de uma população: mutação, migração, deriva genética e seleção natural. Neste capítulo, vamos examinar algumas formas específicas com as quais estas forças moldam a composição genética de populações e produzem a evolução em longo prazo. Começamos examinando como o processo de evolução depende da variação genética e como a variação genética é estudada nas populações naturais. Então voltamos para as mudanças evolutivas que produzem a aparência da nova espécie e como as histórias evolutivas (filogenias) são construídas. Terminamos o capítulo examinando os padrões da mudança evolutiva em nível molecular.

26.1 A evolução ocorre por mudança genética nas populações

A evolução é um dos princípios fundamentais de toda a biologia. Theodosius Dobzhansky, um importante líder dos primóridios da genética evolutiva, uma vez destacou que "nada na biologia faz sentido exceto sob a óptica da evolução". De fato, a evolução é uma teoria abrangente que ajuda a fazer sentido boa parte do mundo natural, desde as sequências de DNA encontradas em nossas células aos tipos de organismos que nos circundam. As evidências da evolução são avassaladoras. A evolução é observada diretamente várias vezes, por exemplo, centenas de diferentes espécies de insetos desenvolveram resistência aos pesticidas comuns que foram introduzidos após a Segunda Guerra Mundial. A evolução é apoiada pelo registro fóssil, pela anatomia comparativa, pela embriologia, pela distribuição de plantas e animais (biogeografia) e pela genética molecular.

Evolução biológica. Apesar da sua grande importância em todos os campos da biologia, a evolução é, com frequência, mal compreendida e mal interpretada. Na nossa sociedade, o termo *evolução* frequentemente é aplicado a qualquer tipo de mudança. Entretanto, a **evolução** biológica refere-se apenas a um tipo específico de mudança, a mudança genética que ocorre em um grupo de organismos. Dois aspectos desta definição devem ser enfatizados. Primeiro, a evolução inclui apenas mudança genética. Muitas mudanças não genéticas ocorrem nos organismos vivos, como o desenvolvimento de uma pessoa inteligente complexa a partir do zigoto unicelular. Embora seja incrível, esta mudança não é evolução, porque não inclui mudanças nos genes. O segundo aspecto enfatiza que a evolução ocorre em *grupos* de organismos. Um organismo não evolui, o que evolui é o *pool* de genes comum a um grupo de organismos.

Evolução como um processo de duas etapas. A evolução pode ser compreendida como um processo de duas etapas. Primeiro, surge a variação genética. A variação genética tem sua origem no processo de mutação, que produz novos alelos e a recombinação, que mistura os alelos em novas combinações. Esses dois processos são aleatórios e produzem variação genética de forma contínua, independente da necessidade da evolução. A segunda etapa no processo de evolução é o aumento e a redução das frequências das variantes genéticas. As várias forças evolutivas discutidas no Capítulo 25 fazem com que alguns alelos no *pool* de genes tenham maior frequência e outros tenham menor frequência. Esta troca na composição do *pool* de genes comum a um grupo de organismos constitui a mudança evolutiva.

Tipos de evolução. Podemos diferenciar entre dois tipos de evolução que ocorrem em um grupo de organismos conectados pela reprodução. A **anagênese** refere-se à evolução que ocorre em um único grupo (uma linhagem) com o passar do tempo (**Figura 26.1**). Outro tipo de evolução é a **cladogênese**, a divisão de uma linhagem em duas. Quando uma linhagem se divide, os dois ramos não têm mais um *pool* de genes em comum e evoluem independentemente um do outro. Novas espécies surgem a partir da cladogênese. Resolva o Problema 21

Conceitos

A evolução biológica é a mudança genética que ocorre em um grupo de organismos. A anagênese é a evolução que ocorre em uma única linhagem, enquanto a cladogênese é a divisão de uma linhagem em duas.

Checagem dos conceitos 1

Descreva resumidamente como a evolução ocorre como um processo de duas etapas.

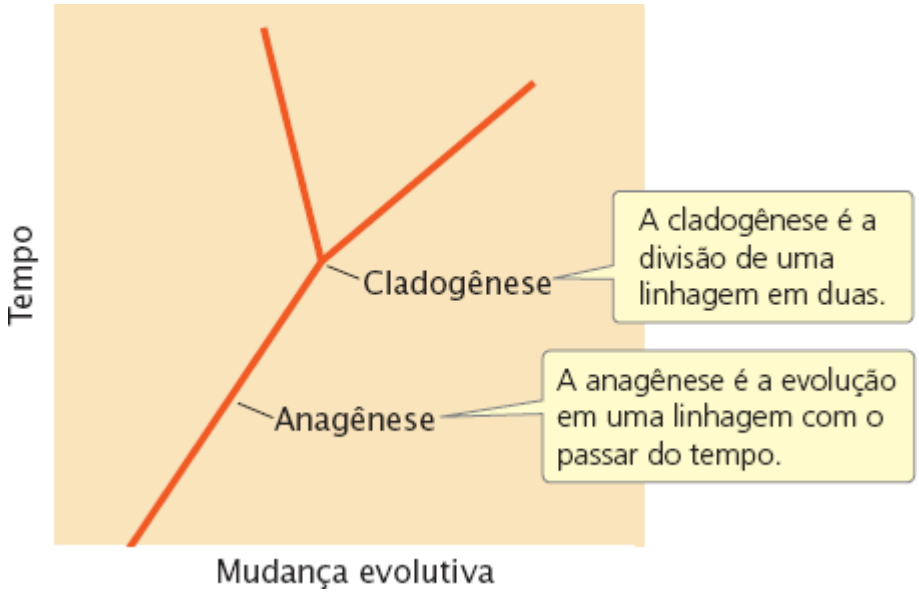


Figura 26.1 A anagênese e a cladogênese são dois tipos diferentes de mudança evolutiva. Anagênese é a mudança em uma linhagem evolutiva, enquanto cladogênese é a divisão das linhagens.

26.2 Muitas populações naturais apresentam níveis elevados de variação genética

Como é obrigatória a variação genética para que ocorra evolução, os biólogos evolutivos estão interessados na variação genética que ocorre nas populações naturais e nas forças que controlam a magnitude e a natureza dessa variação. Por muitos anos, eles não podiam examinar os genes diretamente e eram limitados a estudar os fenótipos dos organismos. Embora a variação genética não possa ser diretamente quantificada, os estudos dos fenótipos sugeriram que muitas populações de organismos abrigam considerável variação genética. As populações de organismos na natureza apresentam variação genotípica significativa; os sapos, por exemplo, apresentam variados padrões de cores, os pássaros têm diferentes tamanhos, as borboletas têm diferentes padrões de manchas, a cor do pelo dos camundongos difere e os humanos apresentam variação no tipo sanguíneo, só para mencionar algumas. Os cruzamentos revelaram que alguns destes traços eram herdados como traços genéticos únicos (**Figura 26.2**), mas, para a maioria dos traços, a base genética precisa era complexa e desconhecida, impedindo que os primeiros geneticistas evolutivos quantificassem a variação genética nas populações naturais.

Conforme discutido no Capítulo 25, a resposta de uma população à seleção depende da herdabilidade de sentido restrito – uma medida da variação genética aditiva de um traço dentro de uma população. Muitos organismos respondem à seleção artificial feita pelos humanos, sugerindo que as populações destes organismos têm muita variação genética aditiva. Por exemplo, os seres humanos usam a seleção artificial para produzir numerosas raças de cães, com incrível variação em tamanho, formato, cor e comportamento (ver Figura 24.21). Estudos iniciais de variações cromossômicas na *Drosophila* e nas plantas também sugeriram que a variação genética nas populações naturais é abundante e comum.

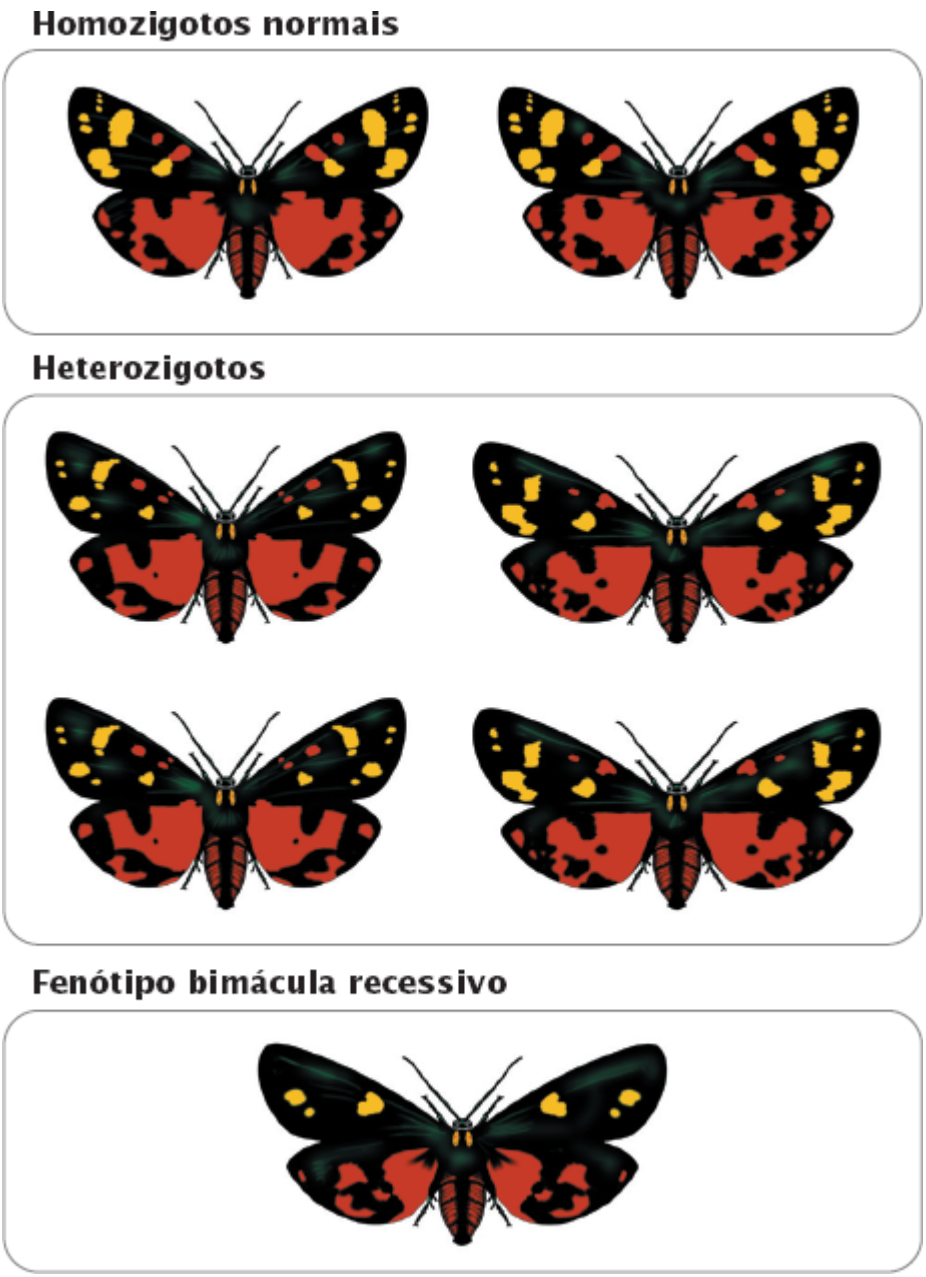


Figura 26.2 Os primeiros geneticistas evolutivos eram forçados a confiar nos traços fenotípicos que tinham uma única base genética. A variação nos padrões de manchas na borboleta *Panaxia dominula* é um exemplo.

Variação molecular

Os avanços na genética molecular possibilitaram investigar a mudança evolutiva diretamente por meio da análise das sequências de proteínas e ácidos nucleicos. Estes dados moleculares oferecem várias vantagens para estudar o processo e o padrão de evolução:

Os dados moleculares são genéticos. A evolução é o resultado da mudança genética com o tempo. Muitos traços anatômicos, comportamentais e psicológicos têm uma base molecular, mas a relação entre os genes envolvidos e o traço é complexa. A variação de sequência de proteínas e ácidos nucleicos tem uma base genética de fácil interpretação.

Os métodos moleculares podem ser usados em todos os organismos. Estudos iniciais de genética de populações dependiam de traços genéticos simples como o tipo sanguíneo humano, padrões de faixas nas cobras ou padrões de manchas nas borboletas (ver Figura 26.2), que estão restritos a um pequeno grupo de organismos. Entretanto, todos os organismos vivos têm proteínas e ácidos nucleicos, então dados moleculares podem ser coletados de qualquer organismo.

Os métodos moleculares podem ser aplicados a uma imensa quantidade de variação genética. Uma gama enorme de dados pode ser acessada pelos métodos moleculares. O genoma humano, por exemplo, tem mais de 3 bilhões de pares de bases de DNA, que constitui um grande *pool* de informações sobre nossa evolução.

Todos os organismos podem ser comparados com o uso de dados moleculares. É difícil tentar acessar a história evolutiva

de organismos distantes porque eles têm poucas características em comum. As relações evolutivas entre os angiospermas eram avaliadas ao comparar a anatomia floral, enquanto as relações evolutivas das bactérias eram determinadas por suas propriedades nutricionais e de coloração. Como as plantas e bactérias têm tão poucas características estruturais em comum, era difícil avaliar como elas estavam relacionadas uma com a outra no passado. Todos os organismos têm alguns traços moleculares em comum, como as sequências de RNA ribossômico e algumas proteínas fundamentais. Estas moléculas oferecem uma base válida para comparações entre todos os organismos.

Os dados moleculares são quantificáveis. Os dados de sequências de proteínas e ácidos nucleicos são precisos, exatos e quantificáveis, o que facilita a avaliação objetiva das relações evolutivas.

Os dados moleculares oferecem, com frequência, informações sobre o processo de evolução. Os dados moleculares podem revelar indícios importantes sobre o processo de evolução. Por exemplo, os resultados de um estudo de sequências de DNA revelaram que um tipo de resistência a inseticida nos mosquitos provavelmente surgiu a partir de uma única mutação que se disseminou pelo mundo.

O banco de dados das informações moleculares é grande e crescente. Atualmente, este banco de dados de DNA e sequências de proteínas pode ser usado para fazer comparações evolutivas e deduzir mecanismos de evolução.

Conceitos

Técnicas moleculares e dados oferecem várias vantagens para os estudos evolutivos. Os dados moleculares são genéticos por natureza e podem ser investigados em todos os organismos; eles fornecem conjuntos de dados potencialmente grandes, permitem que todos os organismos sejam comparados ao usar as mesmas características, são quantificados com facilidade e fornecem informações sobre o processo de evolução.

Variação de proteína

O avanço inicial para quantificar a variação genética nas populações naturais veio com a aplicação de eletroforese (ver Figura 19.2) nos estudos populacionais. Esta técnica separa macromoléculas, como as proteínas e ácidos nucleicos, com base no tamanho e carga. Em 1966, Richard Lewontin e John Hubby extraíram proteínas das moscas-da-fruta, separaram as proteínas por eletroforese e então as coraram para enzimas específicas. Um exame do padrão das bandas no gel permitiu que eles indicassem os genótipos paras moscas individuais e quantificassem a variação genética nas populações naturais. No mesmo ano, Harry Harris quantificou a variação genética nas populações humanas usando a mesma técnica.

Atualmente, numerosos estudos examinam a variação de proteína em uma grande variedade de organismos e descobriram que a maioria das populações tem uma grande quantidade de variação nas proteínas (**Figura 26.3**). Entretanto, a eletroforese de proteína realmente subestima a variação genética, pois ela não detecta algumas substituições de aminoácidos, nem detecta a variação genética no DNA que não altera os aminoácidos de uma proteína (códons sinônimos e variação em regiões não codificadas do DNA).

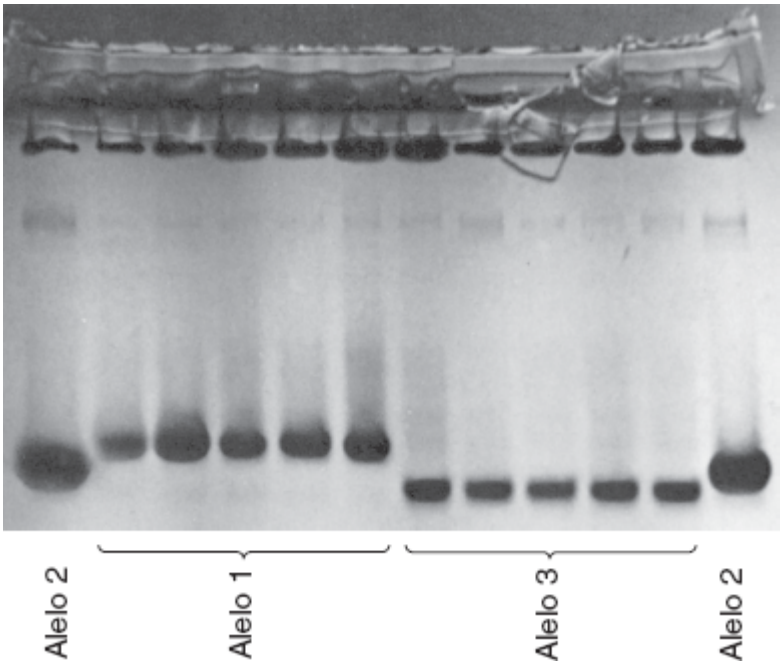


Figura 26.3 A variação molecular nas proteínas é revelada por eletroforese. Amostras de tecido de *Drosophila pseudoobscura* foram submetidas a eletroforese e coradas para esterase. As esterases codificadas por diferentes alelos migram diferentes distâncias. A imagem apresenta no gel os homozigotos para três diferentes alelos. (Cortesia de Richard Lewontin.)

Uma hipótese para explicar a significativa variação molecular nas proteínas é a **hipótese de mutação neutra**, que propõe que parte da variação molecular é adaptativamente neutra, os indivíduos com diferentes variantes moleculares têm aptidão (*fitness*) igual. Essa hipótese, elaborada por Motoo Kimura, não propõe que as proteínas são inúteis, ela sugere que a maioria das variantes tem equivalência funcional. Como estas variantes têm equivalência funcional, a seleção natural não as diferencia e sua evolução é moldada em grande parte pelos processos aleatórios de deriva genética e mutação. A hipótese de mutação neutra aceita que a seleção natural é uma força importante na evolução, mas vê a seleção como um processo que favorece o "melhor" alelo enquanto elimina os outros. Ela propõe que, quando a seleção é importante, existirá *pouca* variação genética.

Outra visão, chamada de **hipótese de equilíbrio**, é que a variação genética nas populações naturais é mantida pela seleção que favorece a variação (seleção de equilíbrio). A sobredominância, na qual o heterozigoto tem maior adaptação do que o homozigoto, é um tipo de seleção de equilíbrio. De acordo com esta hipótese, as variantes moleculares não são fisiologicamente equivalentes e não têm a mesma aptidão. Ao contrário, a variação genética dentro de populações naturais é moldada em grande parte pela seleção e quando a seleção é importante, existirá *muita* variação.

Muitas tentativas para refutar uma hipótese ou outra falharam, porque não estava muito claro o quanto de variação estava de fato presente (lembre-se de que a eletroforese da proteína detecta apenas parte da variação genética) e ambas hipóteses são capazes de explicar muitos padrões diferentes de variação genética. Os resultados de estudos recentes que fornecem informações diretas sobre a variação da sequência de DNA demonstram que parte da variação em nível de DNA tem pouco efeito óbvio sobre o fenótipo; portanto, é provável que seja neutra.

Conceitos

A aplicação da eletroforese no estudo da variação de proteína em populações naturais revelou que a maioria dos organismos apresenta substancial variação genética. A hipótese de mutação neutra propõe que a maioria da variação molecular é neutra com relação a seleção natural e é moldada em grande parte pela mutação e deriva genética. A hipótese de equilíbrio propõe que a variação genética é mantida pela seleção de equilíbrio.

Checagem dos conceitos 2

Qual afirmação sobre a hipótese de mutação neutra é verdadeira?

a. Todas as proteínas são inúteis.
b. A seleção natural não participa da evolução.
c. A maioria das variantes moleculares são funcionalmente equivalentes.
d. Todas as opções acima.

Variação na sequência de DNA

O desenvolvimento de técnicas para isolar, cortar e sequenciar o DNA nos últimos 30 anos forneceu ferramentas úteis para detectar, quantificar e investigar a variação genética. A aplicação destas técnicas permitiu uma visão detalhada da variação genética em nível molecular.

Variação do sítio de restrição. Entre as primeiras técnicas para detectar e analisar a variação genética nas sequências de DNA estava o uso de enzimas de restrição. Cada enzima de restrição reconhece e "recorta" uma sequência específica de nucleotídios de DNA, conhecida como sítio de restrição da enzima (ver *Capítulo 19*). A variação do sítio de restrição é chamada de polimorfismo de comprimento do fragmento de restrição (RFLP; ver Figura 19.19). Cada enzima de restrição reconhece um número limitado de sítios de nucleotídios em um trecho específico de DNA, mas, se forem usadas várias enzimas de restrição diferentes e os sítios reconhecidos pelas enzimas forem sequências aleatórias, os RFLPs podem ser usados para estimar a variação no DNA e a proporção de nucleotídios que diferem entre os organismos. Os RFLPs também podem ser usados para analisar a estrutura genética das populações e para avaliar as relações evolutivas entre os organismos. Os RFLPs foram muito usados nos estudos de evolução antes do desenvolvimento de métodos rápidos e baratos para sequenciar diretamente o DNA e a análise de restrição ainda é usada hoje em dia em alguns estudos de evolução molecular.

Variação de microssatélites. Os microssatélites são sequências curtas de DNA presentes em múltiplas cópias repetidas em *tandem* (ver *Capítulo 19*). A variação no número de cópias das repetições é comum, e organismos individuais apresentam diferentes números de cópias repetidas. Os microssatélites podem ser detectados ao usar a reação em cadeia da polimerase (PCR). Os pares de *primers* são usados para flanquear uma região de cópias repetidas das sequências. Os fragmentos de DNA que são usados na reação de PCR têm diferentes comprimentos, dependendo do número de repetições em *tandem* presentes. O DNA a partir de um organismo individual com mais repetições produzirá um segmento ampliado mais longo. Após a PCR ser completada, os fragmentos ampliados são separados com o uso de eletroforese a gel e corados, produzindo uma série de bandas em um gel (ver Figura 19.25). Os padrões de bandas resultantes representam diferentes alelos e podem ser usados para quantificar a variação genética, avaliar as relações genéticas entre organismos individuais e quantificar as diferenças genéticas em uma população. Uma vantagem de usar microssatélites é que a reação de PCR pode ser usada em quantidades muito pequenas de DNA e é rápida. Os fragmentos amplificados podem ser marcados com corante fluorescente e detectados por *laser*, permitindo que o processo seja automatizado.

David Coltman e colaboradores usaram a variação de microssatélites para estudar a paternidade do carneiro-das-montanhas (ver introdução do *Capítulo 25*) e mostraram que a caça esportiva de carneiros como prêmios reduziu o peso e o tamanho do chifre dos animais. Foram coletadas amostras de sangue, pelo e tecido da orelha do carneiro-das-montanhas em Ram Mountain em Alberta, Canadá, uma população monitorada desde 1971. Foi extraído DNA das amostras de tecidos e ampliadas com PCR, revelando variação em 20 *loci* microssatélites. Com base na variação de microssatélites, foi descoberta a paternidade de 241 carneiros, e as relações familiares dos carneiros foram concluídas. Usando estas relações familiares e as técnicas genéticas quantitativas descritas no Capítulo 24, os geneticistas conseguiram mostrar que o peso do carneiro e o tamanho do chifre apresentavam alta herdabilidade e exibiam uma forte correlação genética positiva. Os caçadores de troféus acertavam seletivamente os carneiros com chifres grandes, antes que eles fossem capazes de se reproduzir. Esta pressão seletiva produziu uma resposta à seleção: os carneiros desenvolveram chifres menores. Entre 1971 e 2002, o tamanho do chifre na população reduziu em cerca de um quarto. Graças à correlação genética positiva entre o tamanho do chifre e o tamanho do corpo, o tamanho do corpo dos carneiros também está reduzindo. A morte de carneiros troféu com grandes chifres levou a uma redução nos vários traços que são cobiçados pelos caçadores. Esta pesquisa ilustra uma aplicação prática do uso de microssatélites na análise evolutiva.

Variação detectada pelo sequenciamento de DNA. As técnicas para sequenciar o DNA de forma rápida e barata (ver *Capítulo 19*) tornaram este tipo de dados uma ferramenta importante nos estudos populacionais e evolutivos. Os dados da sequência de DNA relevam processos que influenciam a evolução e são inestimáveis para determinar as relações evolutivas de diferentes organismos. O uso de PCR para produzir o DNA usado nas reações de sequenciamento significa que os dados podem ser obtidos a partir de uma amostra muito pequena de DNA.

O sequenciamento de DNA foi usado para reavaliar as relações genéticas entre os elefantes africanos. Alfred Roca e colaboradores coletaram amostras de tecido de 195 elefantes ao atingi-los com dardos semelhante a agulhas que caíam no chão após atingir um elefante, mas retinham um pequeno pedaço de pele. A partir das amostras de pele, os cientistas sequenciaram 1.732 pares de bases de DNA a partir de quatro genes nucleares. Sua análise revelou grandes diferenças genéticas entre os elefantes da floresta e da savana, sugerindo que existe um fluxo limitado de genes (migração) entre estes dois grupos de elefantes. Com base nestes resultados, os cientistas propuseram que existem duas espécies diferentes de elefantes na África.

Outro exemplo do uso dos dados da sequência de DNA para decifrar as relações evolutivas foi o caso incomum de infecção

pelo HIV na consulta de dentista na Flórida. Em julho de 1990, o U.S. Centers for Disease Control and Prevention (CDC) relatou que uma mulher jovem na Flórida se tornou HIV-positiva após um procedimento odontológico invasivo feito por um dentista que tinha AIDS. A mulher não tinha fatores de risco conhecidos para infecção pelo HIV nem contato conhecido com pessoas HIV-positivas. O CDC admitiu que ela poderia ter contraído a infecção do seu dentista. Até 1992, sete dos pacientes do dentista tiveram teste positivo para HIV e este número aumentou para dez.

Todos os pacientes do dentista HIV-positivo haviam sido submetidos a procedimentos odontológicos invasivos como tratamento da canal e extração de dente, no período que o dentista estava infectado. Entre os sete pacientes inicialmente estudados pelo CDC (pacientes A-G, **Quadro 26.1**), dois tinham fatores de risco conhecidos para infecção pelo HIV (uso de droga intravenosa, comportamento homossexual ou relações sexuais com pessoas infectadas com HIV) e um terceiro tinha possíveis fatores de risco não confirmados.

Para determinar se o dentista havia infectado seus pacientes, o CDC fez um estudo da evolução molecular do HIV isolado do dentista e dos pacientes. O HIV sofre rápida evolução, o que torna possível rastrear a via de sua transmissão.

Foram coletadas amostras de sangue do dentista, dos pacientes e de um grupo de 35 pessoas do controle local (outras pessoas infectadas pelo HIV que viviam dentro de 145 km do consultório do dentista, mas sem contato com o dentista). Foi extraído o DNA das células brancas e um fragmento de 680 pb do gene do *envelope* do vírus foi amplificado por PCR. Os fragmentos do dentista, pacientes e controles locais foram sequenciados e comparados.

Quadro 26.1 Pessoas HIV-positivas incluídas no estudo de isolados de HIV de um dentista na Flórida.

			Diferenças médias nas sequências de DNA (%)	
Pessoa	**Gênero**	**Fatores de risco conhecidos**	**De HIV do dentista**	**De HIV dos controles**
Dentista	M	Sim		11,0
Paciente A	F	Não	3,4	10,9
Paciente B	F	Não	4,4	11,2
Paciente C	M	Não	3,4	11,1
Paciente E	F	Não	3,4	10,8
Paciente G	M	Não	4,9	11,8
Paciente D	M	Sim	13,6	13,1
Paciente F	M	Sim	10,7	11,9

Fonte: De C. Y. Ou et al., Science 256:1165-1171, 1992, Table 1.

A divergência entre as sequências virais coletadas do dentista, os sete pacientes e os controles está apresentada no Quadro 26.1. O DNA viral coletado dos pacientes sem fatores de risco conhecidos (pacientes, A, B, C, E e G) tinha diferença do DNA viral do dentista em 3,4 a 4,9%, enquanto o DNA dos controles tinha diferença do dentista em uma média de 11%. As sequências virais coletadas de cinco pacientes (A, B, C, E e G) estavam mais intimamente relacionadas às sequências virais coletadas do dentista do que as sequências virais da população geral, sugerindo fortemente que estes pacientes adquiriram sua infecção pelo HIV do dentista. Os isolados virais dos pacientes D e F (pacientes com fatores de risco confirmados), entretanto, diferiam do isolado do dentista em 10,7 a 13,6%, sugerindo que estes dois pacientes não adquiriram sua infecção do dentista.

Uma análise das relações evolutivas das sequências virais confirmou que o vírus coletado do dentista tinha uma relação evolutiva próxima dos vírus coletados dos pacientes A, B, C, E e G (**Figura 26.4**). Os vírus dos pacientes D e F, com fatores de risco conhecidos, não eram mais semelhantes aos vírus do dentista do que os vírus dos controles locais. Dos três pacientes HIV-positivos adicionais que foram identificados desde 1992, apenas um tem as sequências virais que estão intimamente relacionadas com as do dentista.

Este estudo é um excelente exemplo da relevância dos estudos evolutivos para os problemas do mundo real. Como o dentista infectou seus pacientes durante as consultas ainda é um mistério, mas este caso é incomum. Um estudo de quase 16.000 pacientes tratados por profissionais de saúde HIV-positivos não encontrou um único caso de transmissão de HIV confirmado do profissional de saúde para o paciente.

Conceitos

A variação na sequência de nucleotídios do DNA pode ser analisada ao usar os polimorfismos de comprimento do fragmento de restrição, microssatélites e dados de sequenciamento direto do DNA.

Checagem dos conceitos 3

Quais são algumas das vantagens em usar os microssatélites para os estudos evolutivos?

Experimento

Pergunta: O dentista transmitiu HIV a seus pacientes?

Métodos Foram coletadas amostras de vírus do dentista, dos pacientes e de controles locais. Os fragmentos de genes foram sequenciados e comparados.

Resultados

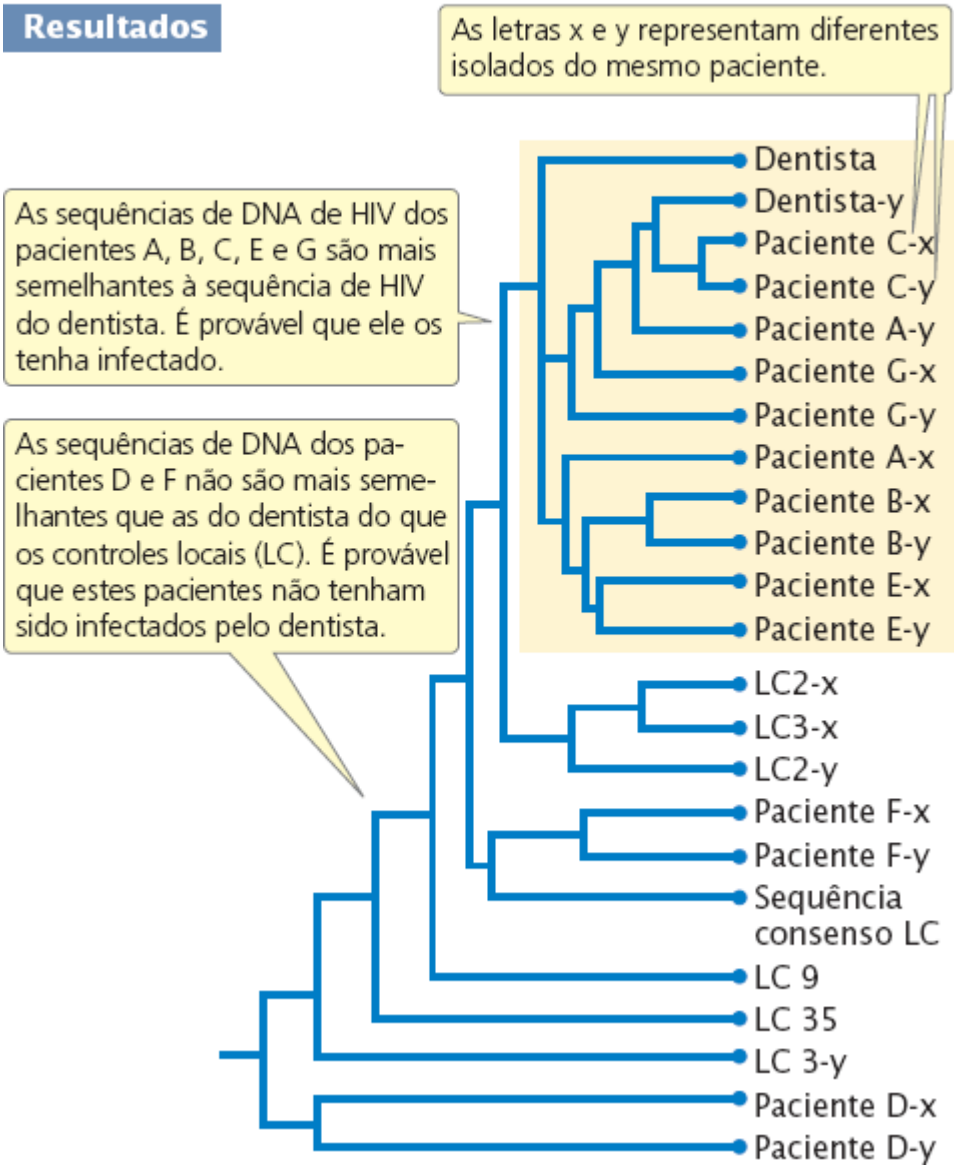


Figura 26.4 Árvore evolutiva mostrando as relações de isolados de HIV de um dentista, sete de seus pacientes (A a G) e de outras pessoas HIV-positivas da mesma região (controles locais, LC). A filogenia baseia-se nas sequências de DNA retiradas do gene do envelope do vírus. (De C. Ou et al., Molecular epidemiology of HIV transmission in a dental practice, Science 256:1167, 1992.)

Conclusão: As sequências virais dos pacientes A, B, C, E e G se agrupam com as do dentista, indicando

uma íntima relação evolutiva. As sequências dos pacientes D e F, junto com os controles locais, estão mais distantes.

26.3 Novas espécies surgem pela evolução do isolamento reprodutivo

Anteriormente, discutimos dois tipos de evolução: anagênese (mudança em uma linhagem) e cladogênese, a divisão das linhagens. A cladogênese surge pela **especiação**, o processo pelo qual uma população se separa em dois grupos evolutivos distintos.

O termo *espécie* significa literalmente tipo ou aspecto; as **espécies** são diferentes variedades ou tipos de organismos vivos. Em muitos casos, é fácil distinguir diferentes espécies: um cavalo é uma espécie diferente de uma galinha. Em alguns casos, entretanto, as diferenças das espécies não são tão claras. Algumas espécies de salamandras *Plethodon* têm aspecto tão semelhante que só podem ser diferenciadas pelo exame de suas proteínas ou genes.

O conceito de espécie tem dois usos principais na biologia. Primeiro, a espécie é um nome dado a um tipo específico de organismo. Para a comunicação efetiva, os biólogos precisam usar um conjunto padrão de nomes para os organismos que eles estudam, e é este o objetivo dos nomes das espécies. Quando um geneticista fala sobre fazer cruzamentos com *Drosophila melanogaster*, outros biólogos entendem imediatamente qual organismo foi usado. O segundo uso do termo espécie está em um contexto evolutivo: uma espécie é considerada um grupo de organismos evolutivamente independentes.

Conceito biológico de espécie

Que tipos de diferenças são necessários para considerar que dois organismos são de espécies diferentes? Uma definição muito usada de espécie é o **conceito biológico de espécie**, desenvolvido pela primeira vez pelo biólogo evolutivo Ernst Mayr em 1942. Mayr estava muito interessado nas características biológicas responsáveis por separar os organismos em unidades de evolução independentes. Ele definiu uma espécie como um grupo de organismos cujos membros são capazes de cruzarem um com outro, mas estão reprodutivamente isolados dos membros de outras espécies. Em outras palavras, os mesmos da mesma espécie têm o potencial biológico para trocar genes e os membros de diferentes espécies não podem trocar genes. Como diferentes espécies não trocam genes, cada espécie evolui de forma independente.

Nem todos biólogos aceitam o conceito biológico de espécie, e existem vários problemas associados com esta ideia. Por exemplo, o isolamento reprodutivo, no qual o conceito biológico de espécie se baseia, não pode ser determinado a partir de fósseis e na prática é difícil determinar se até as espécies vivas são biologicamente capazes de trocar genes. Além disso, o conceito biológico de espécie não pode ser aplicado a organismos que se reproduzem de forma assexuada, como as bactérias. Na prática, a maioria das espécies é distinguida com base nas diferenças fenotípicas (em geral anatômicas). Com frequência, os biólogos partem do pressuposto que diferenças fenotípicas representam diferenças genéticas subjacentes; se os fenótipos de dois organismos são muito diferentes, provavelmente eles não podem e não se cruzam na natureza. Por causa destes problemas, alguns biólogos propuseram definições alternativas para uma espécie. Por exemplo, o *conceito de morfoespécie* define uma espécie com base completamente nas semelhanças e diferenças fenotípicas (morfológicas). O *conceito filogenético de espécie* define uma espécie como o menor grupo reconhecível que tem uma história evolutiva única. Aqui, usaremos o conceito biológico de espécie porque é mais usado e baseia-se nas diferenças genéticas.

Mecanismos de isolamento reprodutivo

A chave para as diferenças das espécies de acordo com o conceito biológico de espécie é o isolamento reprodutivo, características biológicas que impedem que os genes sejam trocados entre diferentes espécies. Qualquer fator ou mecanismo biológico que impeça a troca de genes é chamado de **mecanismo de isolamento reprodutivo**.

Mecanismos de isolamento reprodutivo pré-zigótico. Algumas espécies são separadas por **mecanismos de isolamento reprodutivo pré-zigótico**, que impedem que os gametas de duas espécies se unam e formem um zigoto híbrido. No **isolamento ecológico**, os membros de duas espécies não se encontram um com outro e, portanto, não se reproduzem: eles têm nichos ecológicos diferentes, vivendo em diferentes *habitats* e interagindo com o ambiente de formas diferentes. Por exemplo, algumas espécies de pássaros que habitam florestas se alimentam e fazem ninho na copa das árvores, enquanto outras restringem suas atividades ao solo. Como eles nunca entram em contato, estes pássaros estão reprodutivamente isolados um do outro. Outras espécies são separadas por **isolamento comportamental**, diferenças no comportamento que evitam o intercruzamento. Muitos sapos-machos atraem as fêmeas da mesma espécie ao usar um canto único, específico da espécie. Dois sapos muito próximos podem usar o mesmo lago, mas nunca intercruzam, porque as fêmeas são atraídas apenas pelo canto dos machos de sua espécie.

Outro tipo de isolamento reprodutivo pré-zigótico é o **isolamento temporal**, no qual a reprodução ocorre em períodos

diferentes do ano. Algumas espécies de plantas não trocam genes porque elas florescem em períodos diferentes do ano. O **isolamento mecânico** surge a partir de diferenças anatômicas que evitam uma cópula bem-sucedida. Este tipo de isolamento é observado em muitos insetos, nos quais espécies próximas apresentam diferenças nas genitálias masculina e feminina e a cópula é fisicamente impossível. Finalmente, algumas espécies são separadas pelo **isolamento gamético**, no qual ocorre o cruzamento entre indivíduos de diferentes espécies, mas os gametas não formam zigotos. Os gametas dos machos não sobrevivem no sistema genital feminino ou podem não ser atraídos pelos gametas femininos. Em outros casos, os gametas masculino e feminino se encontram, mas são muito incompatíveis para se unirem e formar um zigoto. O isolamento gamético é observado em muitas plantas, em que o pólen de uma espécie não consegue fertilizar os óvulos de outra espécie.

Mecanismos de isolamento reprodutivo pós-zigótico. Outras espécies são separadas por **mecanismos de isolamento reprodutivo pós-zigótico**, nos quais os gametas de duas espécies se fundem e formam um zigoto, mas não existe fluxo de genes entre as duas espécies, seja porque os híbridos resultantes são inviáveis ou estéreis ou porque não ocorre reprodução nas gerações subsequentes.

Se os mecanismos de isolamento reprodutivo pós-zigóticos falham ou ainda não evoluíram, pode ocorrer acasalamento entre dois organismos de duas espécies diferentes, com a formação de um zigoto híbrido com genes das duas espécies diferentes. Em muitos casos, tais espécies ainda estão separadas por **inviabilidade híbrida**, na qual a incompatibilidade entre os genomas das duas espécies impede que o zigoto híbrido se desenvolva. A inviabilidade híbrida é observada em alguns grupos de sapos, nos quais ocorrem o acasalamento entre diferentes espécies e a fertilização, mas os embriões resultantes nunca atingem desenvolvimento pleno.

Outras espécies são separadas por **esterilidade híbrida**, na qual os embriões híbridos completam o desenvolvimento, mas são estéreis, de modo que os genes não são transmitidos entre as espécies. Os burros e os cavalos acasalam e produzem um descendente viável, uma mula, mas a maioria das mulas é estéril, assim não existe fluxo de genes entre burros e cavalos (veja, porém, o Problema 47, no final do *Capítulo 8*). Finalmente, algumas espécies próximas conseguem acasalar e produzir descendentes F_1 viáveis e férteis. Entretanto, os genes não fluem entre as duas espécies por causa de **colapso do híbrido**, na qual o cruzamento de híbridos produz descendentes inviáveis ou estéreis. O **Quadro 26.2** resume os diferentes tipos de mecanismos de isolamento reprodutivo.

Conceitos

O conceito biológico de espécie a define como um grupo de organismos com potencial de intercruzamento que estão reprodutivamente isolados dos membros de outras espécies. De acordo com este conceito, as espécies estão separadas por mecanismos de isolamento reprodutivo pré-zigótico ou pós-zigótico.

Checagem dos conceitos 4

Qual afirmação é um exemplo de isolamento reprodutivo pós-zigótico?

a. Espermatozoide da espécie A morre no oviduto da espécie B antes de ocorrer a fecundação.
b. Zigotos híbridos entre as espécies A e B são espontaneamente abortados no início do desenvolvimento.
c. As épocas de acasalamento das espécies A e B são diferentes.
d. Os machos das espécies A não são atraídos pelos feromônios produzidos pelas fêmeas da espécie B.

Quadro 26.2 Tipos de mecanismos de isolamento reprodutivo.

Tipo	Características
Pré-zigótico	**Antes de um zigoto ter se formado**
Ecológico	Diferenças no *habitat*, indivíduos não se encontram
Temporal	A reprodução ocorre em períodos diferentes
Mecânico	Diferenças anatômicas impedem a cópula
Comportamental	Diferenças no comportamento do acasalamento evitam o acasalamento
Gamético	Gametas incompatíveis ou não atraídos um por outro
Pós-zigótico	**Após um zigoto ter se formado**

Inviabilidade híbrida	Zigoto híbrido não sobrevive à reprodução
Esterilidade híbrida	O híbrido é estéril
Colapso do híbrido	Os híbridos F_1 são viáveis e férteis, mas os F_2 são inviáveis ou estéreis

Modos de especiação

A **especiação** é o processo no qual surge uma nova espécie. Em relação ao conceito biológico de espécie, a especiação acontece com a evolução dos mecanismos de isolamento reprodutivo.

Existem duas vias principais pelas quais surgem novas espécies. A **especiação alopátrica** surge quando uma barreira geográfica divide uma população em dois grupos e bloqueia a troca de genes entre eles. A interrupção do fluxo de genes leva à evolução de diferenças genéticas que resultam no isolamento reprodutivo. A **especiação simpátrica** é a especiação que surge na ausência de barreira externa para o fluxo de genes; os mecanismos de isolamento reprodutivo se desenvolvem em uma única população. Vamos examinar com mais detalhes estes mecanismos adiante.

Especiação alopátrica. A especiação alopátrica é iniciada quando uma barreira geográfica divide uma população em dois ou mais grupos e impede o fluxo de genes entre os grupos isolados (**Figura 26.5 A**). As barreiras geográficas podem tomar várias formas. O surgimento de uma cadeia de montanhas pode dividir uma população de plantas da baixada em grupos separados em cada lado das montanhas. Os oceanos são barreiras efetivas para muitos tipos de organismos terrestres, separando indivíduos em diferentes ilhas entre si e dos indivíduos do continente. Os rios separam populações de peixes localizadas em drenagens separadas. A erosão das montanhas pode deixar as populações das plantas montanhosas isoladas em picos de montanhas separados.

Após duas populações serem separadas por uma barreira geográfica que evita o fluxo de genes entre elas, elas evoluem de forma independente (**Figura 26.5 B**). O isolamento genético permite que cada população acumule diferenças genéticas que não são encontradas na outra população; as diferenças genéticas surgem pela seleção natural, mutações únicas e deriva genética (se as populações forem pequenas). A diferenciação genética pode levar a isolamento pré-zigótico e pós-zigótico. É importante observar que o isolamento pré-zigótico e o isolamento pós-zigótico surgem simplesmente como consequência de divergência genética.

Se a barreira geográfica que separou as duas populações desaparecer ou os indivíduos conseguirem se dispersar, as populações entram em contato secundário (**Figura 26.5 C**). Neste ponto, vários desfechos são possíveis. Se ocorrer limitada diferenciação genética durante a separação das populações, os mecanismos de isolamento reprodutivo podem não ter evoluído ou podem estar incompletos. Os genes fluirão entre as duas populações, eliminando qualquer diferença genética que tenha surgido e as populações permanecerão como uma única espécie.

Um segundo desfecho possível é que a diferenciação genética durante a separação leva a mecanismos de isolamento reprodutivo pré-zigótico; neste caso, as duas populações são espécies diferentes. Um terceiro desfecho possível é que, durante o tempo de separação, ocorreu alguma diferenciação genética entre as populações, levando a incompatibilidade nos seus genomas e isolamento pós-zigótico. Se os mecanismos de isolamento pós-zigóticos evoluíram, qualquer acasalamento entre indivíduos de diferentes populações produzirá descendentes híbridos inviáveis ou estéreis. Os indivíduos que acasalam apenas com membros da mesma população terão maior aptidão do que os indivíduos que acasalam com membros de outra população; então a seleção natural aumentará a frequência de qualquer traço que evite o cruzamento entre os membros de diferentes populações. Com o passar do tempo, os mecanismos de isolamento reprodutivo pré-zigótico vão evoluir. Em resumo, se existe isolamento reprodutivo pós-zigótico, a seleção natural favorecerá a evolução dos mecanismos de isolamento reprodutivo pré-zigótico para evitar a reprodução desperdiçada de indivíduos acasalando com membros de outra população. Este processo de isolamento reprodutivo pós-zigótico que leva à evolução dos mecanismos de isolamento pré-zigótico é chamado de *reforço*.

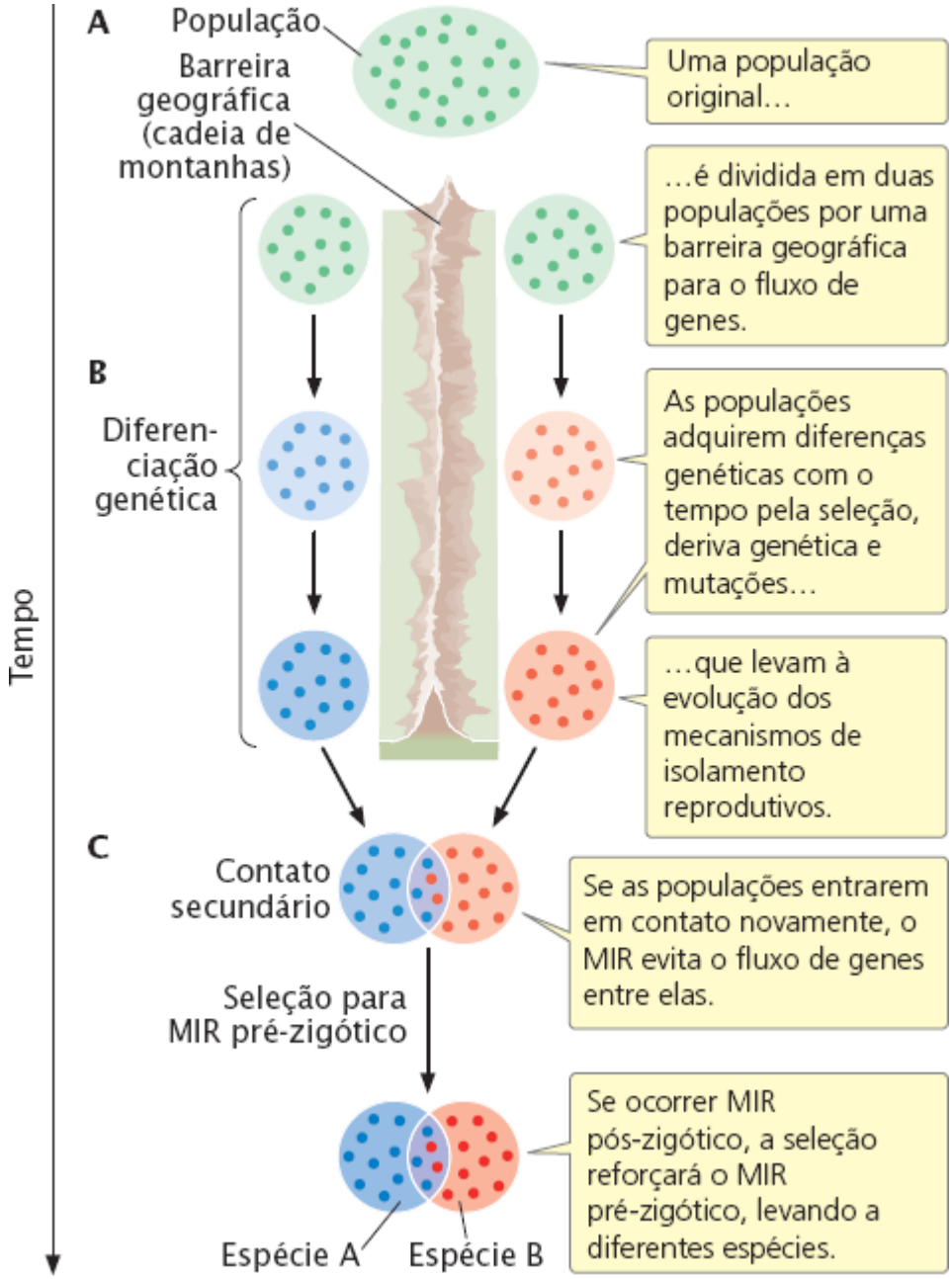


Figura 26.5 A especiação alopátrica é iniciada por uma barreira geográfica para o fluxo de genes entre duas populações. MIR, mecanismo de isolamento reprodutivo.

Figura 26.6 As ilhas Galápagos são jovens do ponto de vista geológico e de origem vulcânica. As ilhas mais antigas estão a leste. (De Philosophical Transactions at the Royal Society of London, Series B 351: 756-772, 1996.)

São possíveis muitas variações neste modelo geral de especiação alopátrica. Muitas novas espécies provavelmente surgem quando um pequeno grupo de indivíduos sofre isolamento geográfico da população principal, por exemplo, alguns indivíduos de uma população do continente podem migrar para uma ilha geograficamente isolada. Nesta situação, o efeito do fundador e a deriva genética são importantes na evolução das diferenças genéticas entre as populações.

Um excelente exemplo de especiação alopátrica pode ser encontrado nos tentilhões-de-darwin, um grupo de aves que Charles Darwin descobriu nas ilhas Galápagos durante sua viagem no navio *Beagle*. Galápagos são um arquipélago de ilhas localizados a cerca de 900 km da costa da América do Sul (**Figura 26.6**). As ilhas Galápagos foram formadas a partir de vulcões que entraram em erupção em um ponto geológico que permaneceu estacionário enquanto a placa geológica se movia para leste há 3 milhões de anos. O movimento da placa geológica deslocou as ilhas recém-formadas para leste, então as ilhas a leste (São Cristóvão e Española) são mais antigas que as a oeste (Isabela e Fernandina). Com o passar do tempo, o número de ilhas do arquipélago aumentou à medida que novos vulcões foram surgindo.

Os tentilhões-de-darwin consistem em 14 espécies encontradas em várias ilhas nesse arquipélago (**Figura 26.7**). As aves variam no formato e tamanho dos seus bicos, que são adaptados para comer diferentes tipos de alimentos. Estudos recentes do desenvolvimento de embriões de tentilhão ajudaram a revelar parte dos detalhes moleculares de como as diferenças no formato dos bicos evoluíram (ver *Capítulo 22*). Estudos genéticos demonstraram que todas as aves estão intimamente relacionadas e evoluíram a partir de uma única espécie ancestral que migrou para as ilhas da costa da América do Sul há 2 a 3 milhões de anos. As relações evolutivas entre as 14 espécies, com base nos estudos de dados de microssatélite, estão descritas na Figura 26.7. A maioria das espécies é separada por um mecanismo de isolamento comportamental (sobretudo o canto), mas algumas das espécies podem e hibridizam na natureza.

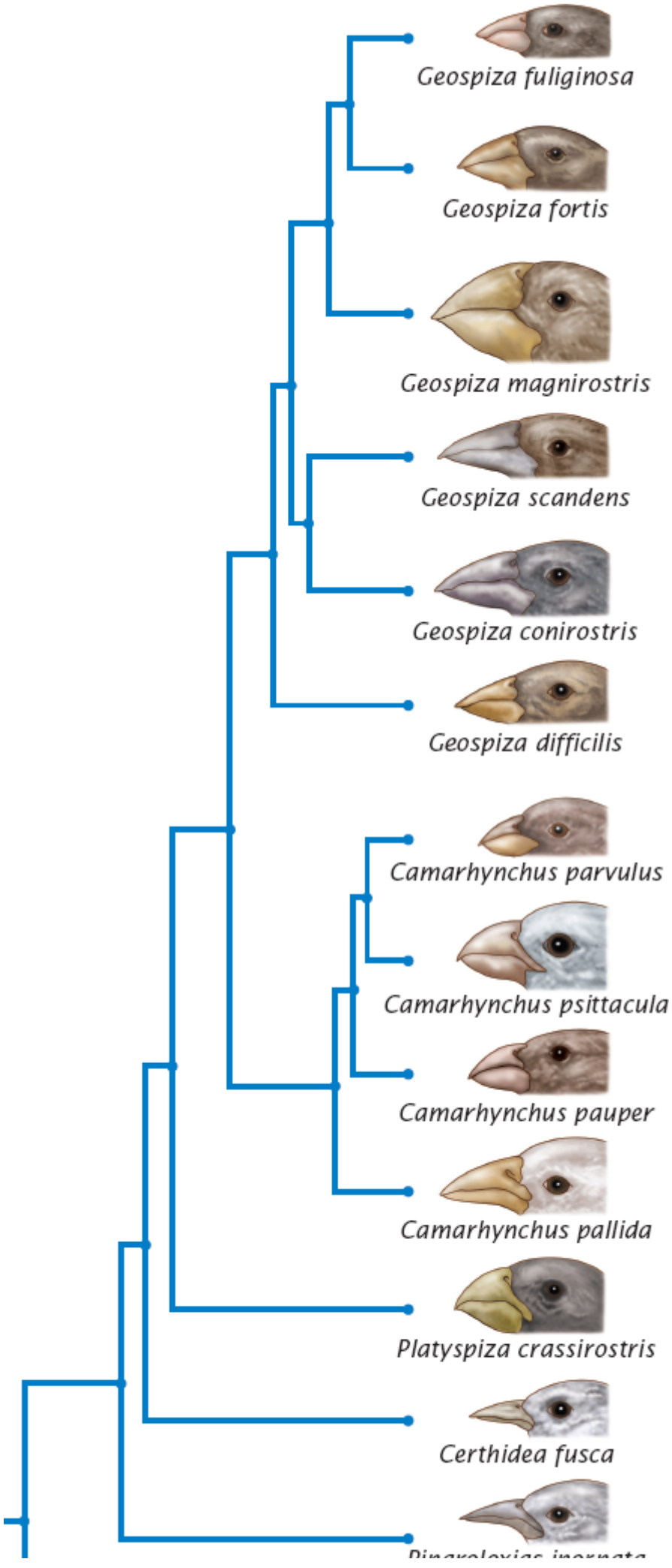

Geospiza fuliginosa
Geospiza fortis
Geospiza magnirostris
Geospiza scandens
Geospiza conirostris
Geospiza difficilis
Camarhynchus parvulus
Camarhynchus psittacula
Camarhynchus pauper
Camarhynchus pallida
Platyspiza crassirostris
Certhidea fusca

Figura 26.7 Os tentilhões-de-darwin consistem em 14 espécies que evoluíram a partir de uma única espécie ancestral que migrou para as ilhas Galápagos e sofreu especiação alopátrica repetida. (De B. R. Grant and P. R. Grant, Bioscience 53:965-975, 2003.)

Os primeiros tentilhões a chegar nas ilhas Galápagos provavelmente colonizaram uma das maiores ilhas orientais. A procriação da população se estabeleceu e aumentou com o tempo. Em algum momento, algumas aves se dispersaram para outras ilhas, onde ficaram isoladas da população original e estabeleceram uma nova população. A nova população sofreu diferenciação genética pela deriva genética e adaptação às condições locais da ilha e consequentemente se tornou isolada do ponto de vista reprodutivo da população original. As aves da nova população então se dispersaram para outras ilhas e deram origem a mais espécies. Este processo foi repetido muitas vezes. Em alguns casos, as aves recém-evoluídas se dispersaram para uma ilha onde existia outra espécie, dando origem a um contato secundário entre as espécies. Atualmente, muitas das ilhas têm mais de um tentilhão residente.

A idade das 14 espécies foi estimada com dados do DNA mitocondrial. A **Figura 26.8** mostra que existe uma forte correspondência entre o número de espécies de aves em vários momentos no passado e o número de ilhas do arquipélago. Esta correspondência é uma das peças mais convincentes para a teoria de que diferentes espécies de tentilhões surgiram a partir da especiação alopátrica.

Conceitos

A especiação alopátrica é iniciada quando uma barreira geográfica do fluxo gênico divide uma única população em duas ou mais populações. Com o passar do tempo, as populações desenvolvem diferenças genéticas, que produzem o isolamento reprodutivo. Após os mecanismos de isolamento reprodutivo pós-zigótico terem evoluído, a seleção favorece a evolução dos mecanismos de isolamento reprodutivo pré-zigótico.

Checagem dos conceitos 5

Qual o papel da deriva genética na especiação alopátrica?

Especiação simpátrica. A especiação simpátrica surge na ausência de qualquer barreira geográfica para o fluxo gênico; os mecanismos de isolamento reprodutivo evoluem em uma única população intercruzada. Há muito tempo a especiação simpátrica é motivo de controvérsia na biologia evolutiva. Ernst Mayr acreditava que a especiação simpátrica era impossível e demonstrou que muitos casos aparentes de especiação simpátrica poderiam ser explicados por especiação alopátrica. Mais recentemente, entretanto, há cada vez mais evidências de que a especiação simpátrica poderia e surge em condições especiais. A dificuldade com a especiação simpátrica é que os mecanismos de isolamento surgem como uma *consequência* da diferenciação genética, que ocorre apenas se o fluxo gênico entre grupos for interrompido. Mas, sem o isolamento reprodutivo (ou alguma barreira externa), como o fluxo gênico pode ser interrompido? Como a diferenciação genética surge em um único grupo que está trocando genes livremente?

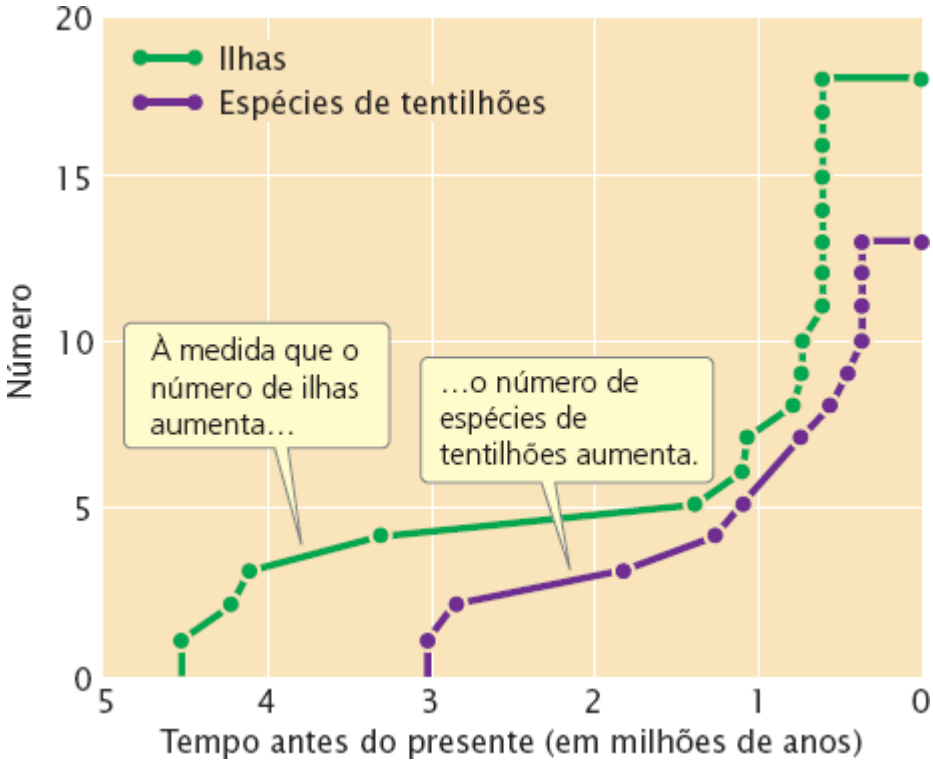


Figura 26.8 O número de espécies dos tentilhões-de-darwin encontradas em vários momentos no passado corresponde ao número de ilhas no arquipélago de Galápagos. (Dados de P. R. Grant, B. R. Grant, and J. C. Deutsch, Speciation and hybridization in island birds, Philosophical Transactions of the Royal Society of London, Series B 351: 765-772, 1996.)

A maioria dos modelos de especiação simpátrica parte do pressuposto que a diferenciação genética é iniciada por uma forte seleção disruptiva que ocorre em uma única população. Um homozigoto (A^1A^1) é fortemente favorecido por um recurso (talvez a espécie de planta que é hospedeira para um inseto) e outro homozigoto (A^2A^2) é favorecido por um diferente recurso (talvez uma diferente planta hospedeira). Os heterozigotos (A^1A^2) têm baixa aptidão em ambos os recursos. Nesta situação, a seleção natural favorecerá os genótipos em outros *loci* que provocam o acasalamento preferencial (acasalamentos entre indivíduos semelhantes, ver *Capítulo 25*), e então não ocorrem acasalamentos entre A^1A^1 e A^2A^2, que produziriam descendentes A^1A^2 com baixa aptidão.

Agora imagine que os alelos em um segundo *locus* afetam o comportamento do acasalamento, de modo que os indivíduos C^1C^1 prefiram acasalar apenas com outros indivíduos C^1C^1 e os indivíduos C^2C^2 prefiram acasalar com outros indivíduos C^2C^2. Se os alelos no *locus A* não estiverem aleatoriamente associados como os alelos no *locus C*, de modo que existam apenas os indivíduos A^1A^1 C^1C^1 e os indivíduos A^2A^2 C^2C^2, então o fluxo gênico será restrito entre os indivíduos usando os diferentes recursos, permitindo que os dois grupos desenvolvam diferenças genéticas que possam levar ao isolamento reprodutivo e à especiação simpátrica.

A dificuldade com este modelo é que a recombinação rompe rapidamente as associações não aleatórias entre genótipos nos dois *loci*, produzindo indivíduos como A^1A^1 C^2C^2, que prefeririam acasalar com os indivíduos A^2A^2 C^2C^2. Este acasalamento produziria todos os descendentes A^1A^2, que teriam poucos recursos. Assim, até a recombinação limitada evitaria a evolução dos genes de preferência do acasalamento.

A especiação simpátrica é mais provável se os genes que afetam a utilização de recurso também afetarem as preferências de acasalamento. Realmente existem genes que afetam a utilização de recursos e a preferência de acasalamento em algumas raças de hospedeiros − populações de insetos especializados que se alimentam em diferentes plantas hospedeiras. Guy Bush estudou o que parecia ser os estágios iniciais de especiação nas raças de hospedeiros da mosca-da-maçã (*Rhagoletis pomonella*). As moscas desta espécie se alimentam de frutas de uma árvore hospedeira específica. O acasalamento ocorre próximo das frutas e as moscas colocam seus ovos nas frutas maduras, onde suas larvas crescem e se desenvolvem. A *R. pomonella* existia originalmente apenas nas frutas dos espinheiros (pilriteiros), que são nativos na América do Norte; mas, 150 anos depois, a *R. pomonella* foi observada pela primeira vez em maçãs cultivadas, que estão relacionadas aos espinheiros, mas uma espécie diferente. As infestações das maçãs por esta nova raça de hospedeiro da *R. pomonella* se espalhou rapidamente e hoje em dia, muitas macieiras na América do Norte estão infestadas por estas moscas.

A raça de hospedeiro da maçã da *R. pomonella* se originou quando novas moscas adquiriram uma mutação que permitiu que elas se alimentassem de maçãs em vez das frutas do espinheiro. Como o acasalamento ocorre sobre as frutas ou perto delas, as moscas que usam maçãs têm maior probabilidade de acasalar com outras moscas que usam as maçãs, levando a um isolamento genético entre as moscas que usam o espinheiro das que usam maçãs. De fato, Bush descobriu que já havia acontecido alguma

diferenciação genética entre as duas raças de hospedeiros. As moscas colocam seus ovos na fruta madura e existe uma forte seleção para as moscas sincronizarem sua reprodução com o período quando sua espécie hospedeira tem fruta madura. As maçãs amadurecem semanas antes do que as frutas dos espinheiros. O período de pico de acasalamento das moscas nas macieiras é 3 semanas antes do que no espinheiro. Estas diferenças na sincronização da reprodução entre a maçã e o fruto do espinheiro reduziu o fluxo gênico em cerca de 2% entre as duas raças hospedeiras e levou a significativa diferenciação genética entre elas. Estas diferenças evoluíram nos últimos 150 anos e parece que a evolução está em andamento. Embora a diferenciação genética ocorra entre as raças hospedeiras de maçã e espinheiro da *R. pomonella* e algum grau de isolamento reprodutivo tenha se desenvolvido entre elas, esse isolamento reprodutivo não está ainda completo e a especiação não aconteceu completamente. **Resolva o Problema 22**

Especiação por poliploidia. Um tipo especial de especiação simpátrica ocorre pela poliploidia (ver *Capítulo 8*). Os organismos poliploides têm mais de dois genomas (3*n*, 4*n*, 5*n* etc.). Como discutido no Capítulo 8, a alopoliploidia surge quando duas espécies diploides hibridizam, produzindo descendentes 2*n* híbridos. A não disjunção em um dos descendentes híbridos produz um tetraploide 4*n*. Como este tetraploide tem exatamente duas cópias de cada cromossomo, em geral é fértil e estará isolado do ponto de vista reprodutivo das duas espécies parentais por diferenças no número de cromossomos (ver Figura 8.26).

Várias espécies de plantas florescentes são alopoliploides. A especiação pela poliploidia foi observada quando ela levou uma nova espécie de morraça que surgiu ao longo da costa da Inglaterra em 1870. Este poliploide tem genomas da morraça europeia *Spartina maritima* (2*n* = 60) e morraça americana *S. alterniflora* (2*n* = 62, **Figura 26.9**).

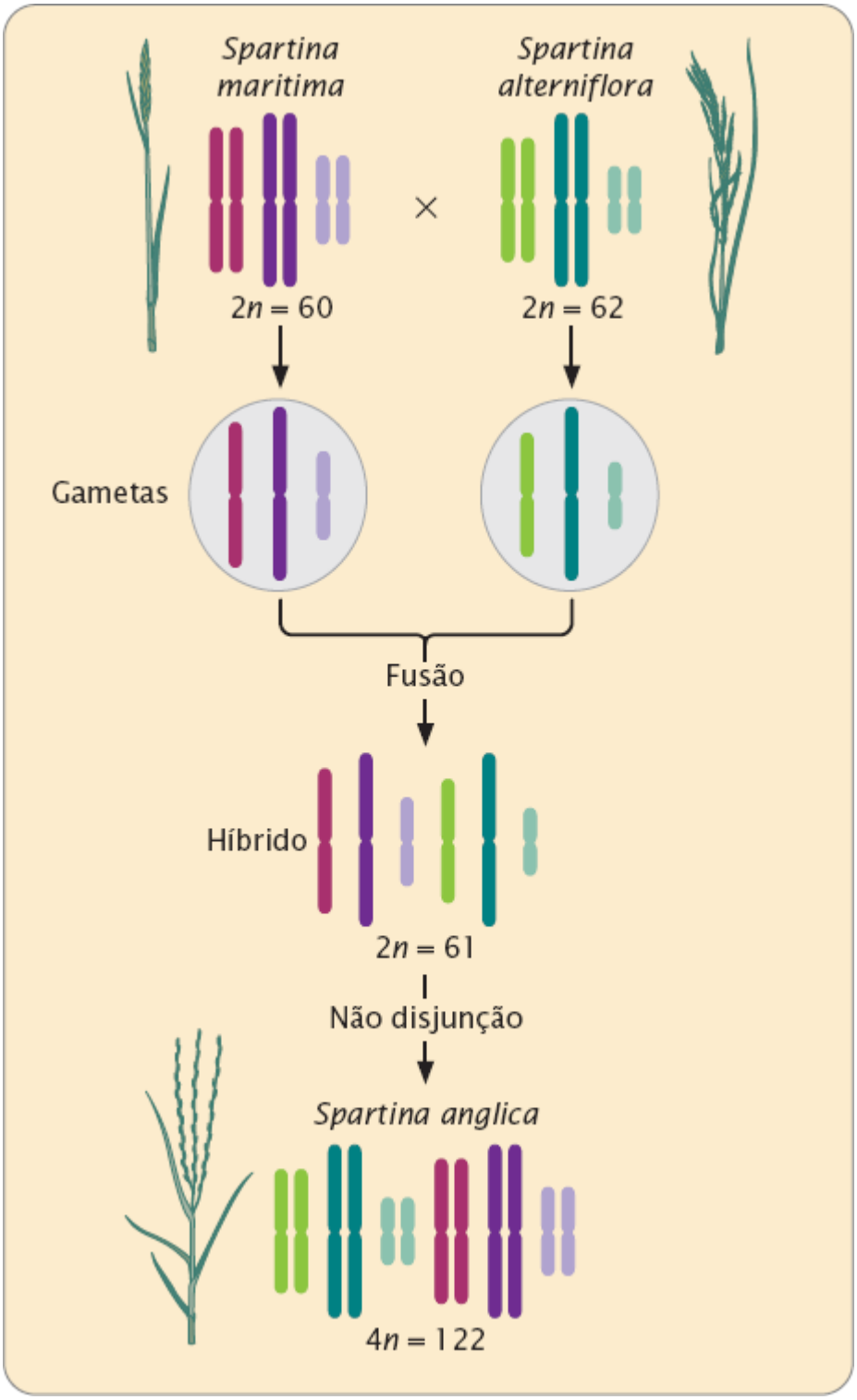


Figura 26.9 ***Spartina anglica*** **surgiu simpatricamente por alopoliploidia.**

As sementes da morraça americana foram transportadas para a Inglaterra no lastro de um navio. Independentemente de como chegou lá, *S. alterniflora* cresceu em um pântano inglês e acabou cruzando com *S. maritima*, produzindo um híbrido com $2n = 61$. A não disjunção no híbrido então levou à dobra dos cromossomos, produzindo uma nova espécie, *S. anglica*, com $4n = 122$ (ver Figura 26.9). Esta nova espécie se espalhou ao longo da costa da Inglaterra.

Conceitos

A especiação simpátrica surge em uma única população com intercruzamento sem qualquer barreira geográfica para o fluxo de genes. A especiação simpátrica pode surgir sob condições especiais, como quando o uso do recurso está ligado à preferência de acasalamento (em raças hospedeiras) ou quando a hibridização da espécie leva a alopoliploidia.

Diferenciação genética associada com a especiação

Como observamos, a diferenciação genética leva à evolução dos mecanismos de isolamento reprodutivo, que restringem o fluxo gênico entre populações e leva à especiação. Quanto de diferenciação genética é necessário para que ocorra o isolamento reprodutivo? Esta pergunta tem sido estudada por geneticistas evolutivos, mas, infelizmente, não existe uma resposta universal. Algumas espécies recém-formadas apresentam vários genes diferentes, enquanto outras parecem ter sofrido divergência em apenas alguns genes.

Um grupo de organismos que foi muito estudado sobre as diferenças genéticas associadas com a especiação é o gênero *Drosophila*. O grupo *Drosophila willistoni* é composto por 12 espécies encontradas nas Américas Central e do Sul em vários estágios no processo de especiação. Usando eletroforese de proteínas, Francisco Ayala e colaboradores obtiveram o genótipo das moscas de diferentes populações geográficas (populações com limitadas diferenças genéticas), subespécies (populações com consideráveis diferenças genéticas), espécies irmãs (espécies recém-surgidas) e espécies não irmãs (espécies mais antigas). Para cada grupo, eles computaram uma medida de similaridade genética, que varia de 1 a 0 e representa o nível geral de diferenciação genética (**Quadro 26.3**). Eles descobriram que existe uma redução geral na similaridade genética à medida que as moscas evoluem de populações geográficas para subespécies de espécies irmãs para espécies não irmãs. Estes dados sugerem que é necessária considerável diferenciação genética em muitos *loci* para que a especiação surja. Um estudo com *D. simulans* e *D. melanogaster*, duas espécies que produzem híbridos inviáveis quando cruzadas, sugeriu que pelo menos 200 genes contribuem para a inviabilidade dos híbridos entre as duas espécies.

Entretanto, outros estudos sugerem que a especiação tenha surgido em decorrência de mudanças em apenas alguns genes. Por exemplo, *D. heteroneura* e *D. silvestris* são duas espécies de moscas-da-fruta havaianas que exibem isolamento reprodutivo comportamental. O isolamento é determinado em grande parte por diferenças no formato da cabeça; *D. heteroneura* tem uma cabeça com formato de martelo com olhos bem separados que são reconhecidos pelas fêmeas da mesma espécie, mas rejeitados pelas fêmeas de *D. silvestris*. Os estudos genéticos indicam que apenas alguns *loci* (cerca de 10) determinam as diferenças no formato da cabeça.

Em outro estudo, os pesquisadores examinaram as diferenças genéticas em duas espécies de aves muito próximas, o papa-mosca-de-colar (*Ficedula albicollis)* e o papa-mosca-preto (*F. hypoleuca*), encontradas na Europa. Acredita-se que essas duas espécies tenham sido separadas a partir de um ancestral comum há menos de 2 milhões de anos, provavelmente por especiação alopátrica associada com a glaciação. Nas áreas onde ambas espécies ocorrem juntas, os papa-moscas-de-colar e os papa-moscas-pretos acasalam, mas as duas espécies são separadas por mecanismos de isolamento reprodutivo pré- e pós-zigótico. Por exemplo, as fêmeas preferem os machos com plumagem e canto da sua própria espécie e os híbridos não sobrevivem tão bem como os descendentes dos genitores de ambas espécies.

Quadro 26.3 Similaridade genética nos grupos do complexo *Drosophila willistoni*.

Grupo	Similaridade genética média
Populações geográficas	0,970
Subespécies	0,795
Espécies irmãs	0,517
Espécies não irmãs	0,352

Pesquisadores da Uppsala University na Suécia sequenciaram completamente os genomas de 10 aves-machos de cada espécie. Eles descobriram que grandes partes dos genomas das duas espécies eram iguais: por exemplo, apenas 5 em cada 1.000 nucleotídios eram diferentes entre as espécies, comparados com 4 de cada 1.000 nucleotídios nas espécies. As diferenças que realmente existiam entre as espécies não eram uniformes ao longo do genoma; em vez disso, as diferenças tendiam a ficar concentradas em algumas “ilhas de divergência” que tinham diferenças de sequências até 50 vezes maiores do que a média para o genoma. Foram encontradas algumas ilhas de divergências em cada cromossomo, e elas ocorriam principalmente perto de centrômeros e telômeros, sugerindo que as diferenças na estrutura do cromossomo são importantes na especiação. As ilhas de divergência tendiam a mostrar baixos níveis de diferenças entre as espécies, indicando que a seleção natural favoreceu diferentes genes em cada região.

Conceitos

Algumas espécies que acabaram de surgir têm um número considerável de diferenças genéticas; outras têm poucas diferenças genéticas.

26.4 A história evolutiva de um grupo de organismos pode ser reconstruída pelo estudo das mudanças nas características homólogas

As relações evolutivas de um grupo de organismos são chamadas de **filogenia**. Como a maior parte da evolução ocorre em longos períodos de tempo e não é fácil a observação direta, os biólogos precisam reconstruir as filogenias ao deduzir as relações evolutivas entre os organismos atuais. A descoberta de fósseis de organismos ancestrais pode auxiliar na reconstrução das filogenias, mas o registro do fóssil não é, com frequência, de grande valia. Assim, os biólogos ficam restritos a análises de características nos organismos atuais para determinar suas relações evolutivas. No passado, as relações filogenéticas eram reconstruídas com base nos traços fenotípicos, frequentemente, traços anatômicos. Atualmente, os dados moleculares, incluindo proteínas e sequências de DNA, são frequentemente usados para construir árvores filogenéticas.

As filogenias são reconstruídas por inferência de mudanças que ocorreram em características homólogas. Tais características evoluíram da mesma característica em um ancestral comum. Por exemplo, embora a perna dianteira de um camundongo e a asa de um morcego tenham aspectos e funções diferentes, o exame minucioso de suas estruturas e de seu desenvolvimento revela que são, na verdade, homólogos; ambos evoluíram do mesmo membro dianteiro de um mamífero anterior, que foi ancestral do camundongo e do morcego. E, como o camundongo e o morcego têm essas características homólogas e outras em comum, sabemos que ambos são mamíferos. Da mesma forma, as sequências de DNA são homólogas se duas sequências atuais evoluíram de uma única sequência encontrada em um ancestral. Por exemplo, todos os organismos eucarióticos têm um gene para o citocromo *c*, uma enzima que participa da respiração oxidativa. Acredita-se que este gene tenha surgido em um único organismo no passado distante e tenha sido transmitido para os descendentes desse ancestral distante. Atualmente, todas as cópias do gene do citocromo *c* são homólogas, porque evoluíram da mesma cópia original no ancestral distante de todos os organismos que tenham este gene.

Uma representação gráfica de uma filogenia é chamada de **árvore filogenética**. Como apresentado na **Figura 26.10**, uma árvore filogenética descreve as relações evolutivas entre diferentes organismos, da mesma forma como um heredograma representa as relações genealógicas entre os membros de uma família. Uma árvore filogenética é composta por **nós (pontos de ramificação)** que representam os diferentes organismos que são comparados, que podem ser indivíduos, populações ou espécies diferentes. Os nós terminais (os que estão nos ramos mais distantes da árvore) representam os organismos para os quais foram obtidos os dados, em geral os organismos atuais. Os nós internos representam os ancestrais comuns que existiram antes que ocorresse a divergência entre os organismos. Na maioria dos casos, os nós internos representam ancestrais anteriores que foram deduzidos a partir da análise. Os nós estão conectados por **ramos**, que representam as conexões evolutivas entre os organismos. Em algumas árvores filogenéticas, os comprimentos dos ramos representam a divergência evolutiva que ocorreu entre os organismos. Quando um nó interno representa um ancestral comum a todos os outros nós na árvore, a árvore é chamada de **árvore filogenética com raiz**. As árvores têm raiz ao incluir na análise um organismo distante de todos os outros; esse organismo distante, mas relacionado, é chamado de grupo externo.

As árvores filogenéticas são criadas para descrever as relações evolutivas entre os organismos; também são criadas para descrever as relações evolutivas entre sequências de DNA. O segundo tipo de árvore filogenética é chamado de **árvore gênica** (**Figura 26.11**). Resolva o Problema 27

Conceitos

Uma filogenia representa as relações evolutivas entre um grupo de organismos e é descrita graficamente por uma árvore filogenética, que é composta por nós que representam os organismos e ramos que representam suas conexões evolutivas.

Checagem dos conceitos 6

Qual característica é encontrada em uma árvore com raiz, mas não em uma árvore sem raiz?

a. Nós terminais.
b. Nós internos.
c. Um ancestral comum a todos os outros nós.
d. Comprimentos dos ramos que representam a divergência evolutiva entre os nós.

Alinhamento de sequências homólogas

Atualmente, as árvores filogenéticas são construídas a partir de dados de sequências de DNA. Esta construção requer que as sequências homólogas sejam comparadas. Assim, a primeira etapa na construção das árvores filogenéticas a partir de dados de sequências de DNA é identificar os genes homólogos e alinhar adequadamente suas bases de nucleotídios. Considere as sequências a seguir que podem ser encontradas em dois organismos diferentes:

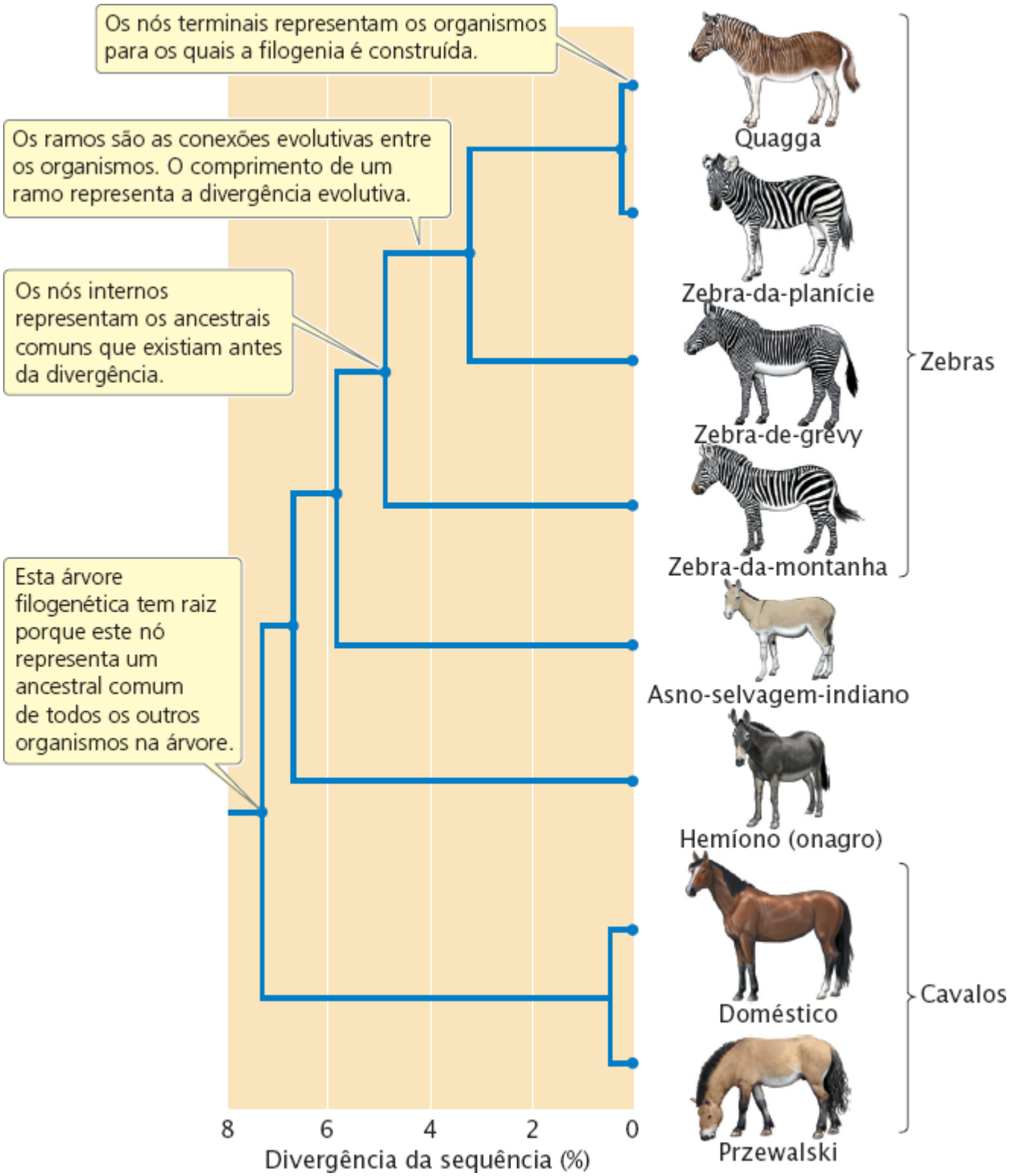


Figura 26.10 **Uma árvore filogenética é uma representação gráfica das relações evolutivas entre um grupo de organismos.**

Posição do nucleotídio	1 2 3 4 5 6 7 8
Gene *X* da espécie A	5′–A T T G C G A A–3′
Gene *X* da espécie B	5′–A T G C C A A C–3′

Essas duas sequências podem ser alinhadas de várias formas possíveis. Podemos assumir que ocorreram substituições de base nas posições 3, 4, 6 e 8:

Posição do nucleotídio	1 2 3 4 5 6 7 8
Gene *X* da espécie A	5′–A T T G C G A A–3′
Gene *X* da espécie B	5′–A T G C C A A C –3′

Por outro lado, podemos assumir que um nucleotídio na posição 3 foi inserido ou eliminado, gerando um espaço na sequência da espécie B e que existe uma substituição de nucleotídio único na posição 6:

Posição do nucleotídio	1 2 3 4 5 6 7 8
Gene *X* da espécie A	5′–A T T G C G A A–3′
Gene *X* da espécie B	5′–A T – G C C A A C –3′

O segundo alinhamento requer menos etapas evolutivas (uma deleção ou inserção mais uma substituição de base) do que o primeiro alinhamento (quatro substituições de base). Os alinhamentos de sequência são feitos por programas de computador que incluem suposições sobre quais tipos de mudanças têm mais probabilidade de ocorrer. Se duas sequências sofreram muita divergência, então pode ser difícil gerar os alinhamentos.

Construção das árvores filogenéticas

Considere uma única filogenia que descreve as relações evolutivas entre três organismos, humanos, chimpanzés e gorilas. Charles Darwin propôs originalmente que os chimpanzés e gorilas estavam intimamente relacionados com os humanos e a pesquisa moderna defende uma relação próxima entre estas três espécies. Existem três possíveis árvores filogenéticas para humanos, chimpanzés e gorilas (**Figura 26.12**). O objetivo do biólogo evolutivo é determinar qual das três árvores está correta. Os dados moleculares aplicados a esta pergunta sugerem uma relação próxima entre humanos e chimpanzés. Para compreender a dificuldade em construir as árvores filogenéticas, vamos considerar por um momento o número de todas árvores possíveis que existem para um grupo de organismos. O número de árvores com raiz possíveis para um grupo de organismos é:

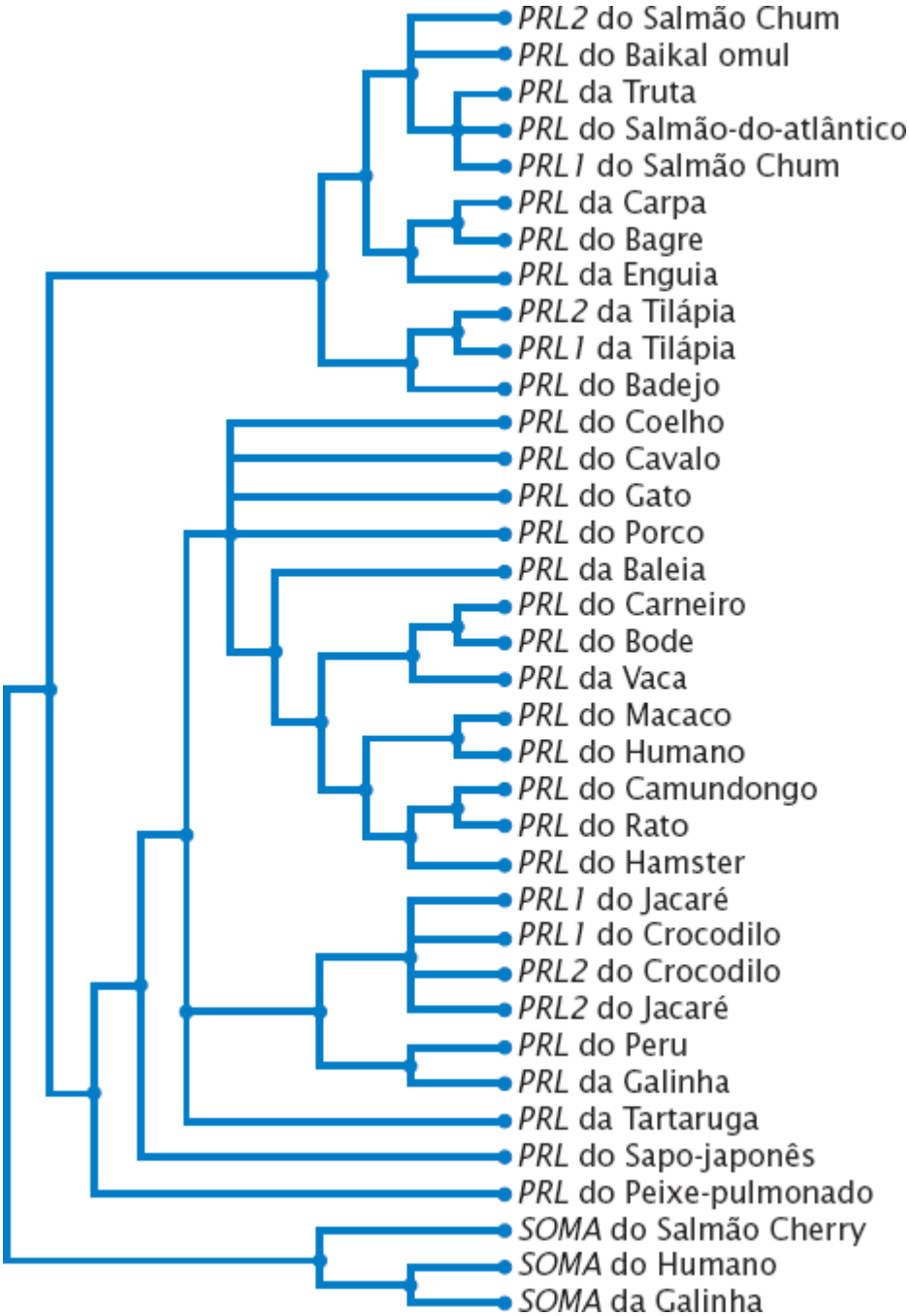


Figura 26.11 Uma árvore genealógica pode ser usada para representar as relações evolutivas entre um grupo de genes. Esta árvore genealógica é uma árvore com raiz, na qual *PRL* representa um gene da prolactina, *PRL1* e *PRL2* são dois genes de prolactina diferentes encontrados no mesmo organismo; e *SOMA* representa um gene da somatropina, que está relacionado com os genes da prolactina. (De M. P. Simmons and J. V. Freudestein, Uninode coding vs. gene tree parsimony for phylogenetic reconstruction using duplicate genes, Molecular Phylogenetics and Evolution 23:488, 2002.)

$$\text{Número de árvores com raiz} = \frac{(2N-3)!}{2^{N-2}(N-2)!}$$

Em que N é igual ao número de organismos incluídos na filogenia e o símbolo ! representa o fatorial, o produto de todos os integrantes de N a 1. Substituindo os valores de N nesta equação, encontramos:

Número de organismos incluídos na filogenia (*N*)	Número de árvores com raiz
2	1
3	3
4	15
5	105
10	34.459.425
20	$8{,}2 \times 10^{21}$

Como o número de organismos na filogenia aumenta além de apenas alguns, o número de árvores com raiz possíveis se torna astronômico. Evidentemente, é impossível escolher a melhor árvore ao comparar todas as possibilidades.

Existem várias abordagens diferentes para deduzir as relações evolutivas e construir as árvores filogenéticas. Em uma abordagem, chamada de *abordagem por distância*, as relações evolutivas são deduzidas com base no grau geral de similaridade entre os organismos. Várias características fenotípicas diferentes ou sequências de genes são examinadas e os organismos são agrupados com base na sua similaridade geral, levando em consideração todas as características e sequências examinadas. Uma segunda abordagem, chamada *abordagem da máxima parcimônia*, deduz as relações filogenéticas com base no menor número de mudanças evolutivas que devem ter ocorrido desde que os organismos tiveram pela última vez um ancestral em comum. Uma terceira abordagem, chamada de *máxima verossimillhança* e *métodos bayesianos*, deduz as relações filogenéticas com base em qual filogenia maximiza a probabilidade de obter o conjunto de características exibidas pelos organismos. Nesta abordagem, uma filogenia com maior probabilidade de produzir as características observadas nos organismos estudados é preferida em relação à filogenia com menor probabilidade.

Com todas as três abordagens para construir filogenias, existem vários métodos numéricos diferentes para a construção das árvores filogenéticas. Todos incluem algumas suposições que ajudam a limitar o número de diferentes árvores que devem ser consideradas; a maioria depende de programas de computador que comparam características fenotípicas ou dados de sequências para agrupar sequencialmente os organismos na construção da árvore.

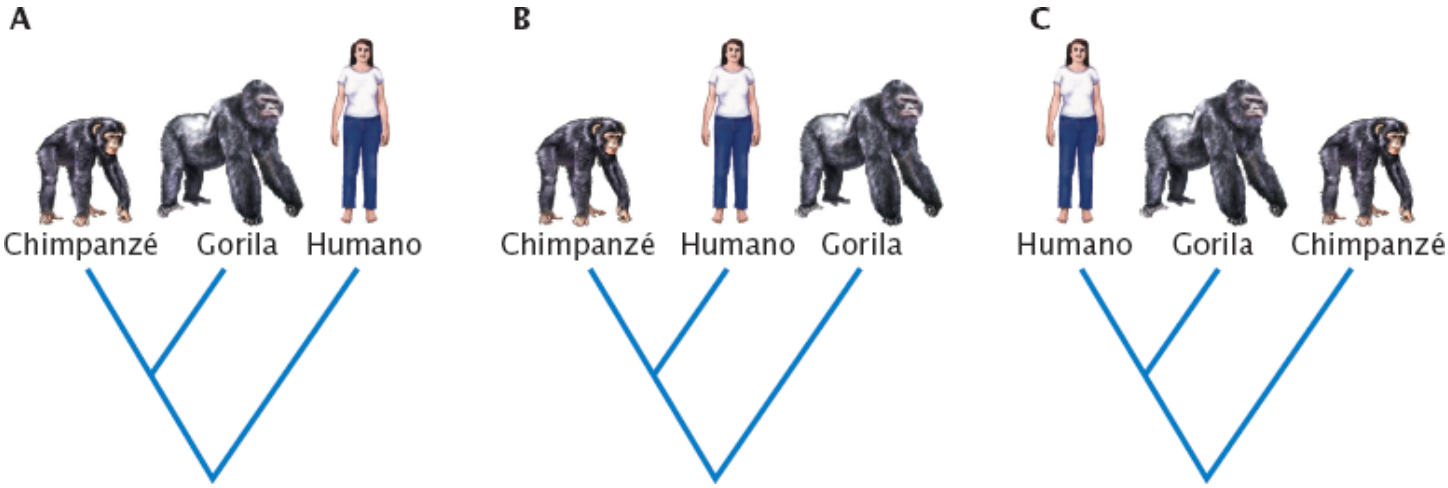


Figura 26.12 **Existem três possíveis árvores filogenéticas para humanos, chimpanzés e gorilas.**

Conceitos

Os dados moleculares podem ser usados para deduzir as filogenias (histórias evolutivas) de grupos de organismos vivos. A construção das filogenias requer o alinhamento adequado de sequências homólogas de DNA. São feitas várias abordagens diferentes para reconstruir as filogenias, incluindo métodos por distância, métodos de máxima parcimônia e máxima verossimilhança e métodos bayesianos.

26.5 Os padrões de evolução são revelados por mudanças moleculares

A capacidade de analisar a variação genética em nível molecular revelou vários processos evolutivos e características que não se suspeitava antes. Esta seção analisa vários aspectos da evolução em nível molecular.

Taxas de evolução molecular

Os achados a partir de estudos moleculares de numerosos genes demonstraram que diferentes genes e diferentes partes do mesmo gene evoluem em diferentes taxas.

Taxas de substituição de nucleotídios. As taxas de mudança evolutiva nas sequências de nucleotídios são medidas como a taxa de substituição de nucleotídios, que é o número de substituições que ocorrem por sítio de nucleotídio por ano dentro de uma população. Para calcular a taxa de substituição de nucleotídios, começamos observando as várias sequências a partir de diferentes organismos. Primeiro alinhamos as sequências homólogas e então comparamos as sequências e determinamos o número de nucleotídios que diferem entre as duas sequências. Podemos comparar as sequências do hormônio do crescimento (GH) para camundongos e ratos, que divergiram de um ancestral comum há 15 milhões de anos. A partir do número de

nucleotídios diferentes nos seus genes do GH, computamos o número de substituições de nucleotídios que devem ter ocorrido desde que divergiram. Como o mesmo sítio pode ter sofrido mutação mais de uma vez, o número de substituições de nucleotídios é maior que o número de diferenças de nucleotídios nas suas sequências e foram desenvolvidos métodos matemáticos especiais para deduzir o número real de prováveis substituições que tenha ocorrido.

Quando temos o número de substituições de nucleotídios por sítio de nucleotídio, dividimos pelo intervalo de tempo evolutivo que separa os dois organismos (em geral obtido a partir de registro fóssil) para obter uma taxa geral de substituição de nucleotídios. Para o gene do hormônio do crescimento do camundongo e rato, a taxa geral de substituição de nucleotídio é de aproximadamente 8×10^{-9} substituições por sítio por ano.

Taxas de substituições não sinônima e sinônima. As mudanças de nucleotídios em um gene que alteram a sequência de aminoácidos de uma proteína são chamadas de substituição não sinônima. As mudanças de nucleotídios, sobretudo os que estão na terceira posição de um códon, que não alteram a sequência de aminoácidos, são chamadas de substituições sinônimas. A taxa de substituições não sinônimas varia muito entre os genes dos mamíferos. A taxa para a proteína α-actina é de apenas $0{,}01 \times 10^{-9}$ substituições por sítio por ano, enquanto a taxa para interferona γ é de $2{,}79 \times 10^{-9}$, quase 300 vezes superior. A taxa de substituição sinônima também varia entre os genes, mas não na magnitude da variação na taxa da não sinônima. Para a maioria dos genes codificadores de proteínas, a taxa de mudança sinônima é consideravelmente maior que a taxa não sinônima porque as mutações sinônimas são toleradas pela seleção natural (**Quadro 26.4**). As mutações não sinônimas, por outro lado, alteram a sequência de aminoácidos da proteína e em muitos casos, são prejudiciais para a aptidão do organismo; parte destas mutações são eliminadas pela seleção natural.

Taxas de substituição para diferentes partes de um gene. Diferentes partes de um gene também evoluem em diferentes taxas, com as maiores taxas de substituições em regiões do gene que têm o menor efeito na função, como a terceira posição de um códon, regiões flanqueadoras e íntrons (**Figura 26.13**). As regiões flanqueadoras 5′ e 3′ dos genes não são transcritas em RNA, portanto, as substituições nestas regiões não alteram a sequência de aminoácidos da proteína, embora possam afetar a expressão gênica (ver *Capítulos 16* e *17*). As taxas de substituição nos íntrons são quase tão altas quanto estas. Embora estes nucleotídios não codifiquem aminoácidos, os íntrons devem ser recompostos no pré-mRNA para ser produzida uma proteína funcional e são necessárias sequências específicas nos locais de recomposição 5′ e 3′ e no ponto de ramificação para a recomposição correta (ver *Capítulo 14*).

Quadro 26.4 Taxas de substituições não sinônimas e sinônimas nos genes de mamíferos com base nas comparações humano–roedor.

Gene	Taxa não sinônima (por sítio por 10^9 anos)	Taxa sinônima (por sítio por 10^9 anos)
α-Actina	0,01	3,68
β-Actina	0,03	3,13
Albumina	0,91	6,63
Aldolase A	0,07	3,59
Apoproteína E	0,98	4,04
Creatinoquinase	0,15	3,08
Eritropoetina	0,72	4,34
α-Globina	0,55	5,14
β-Globina	0,80	3,05
Hormônio do crescimento (GH)	1,23	4,95
Histona 3	0,00	6,38
Cadeia pesada da imunoglobulina (região variável)	1,07	5,66

Insulina	0,13	4,02
Alfainterferona 1	1,41	3,53
Gamainterferona	2,79	8,59
Hormônio luteinizante (LH)	1,02	3,29
Somatostatina-28	0,00	3,97

Fonte: De W. Li and D. Graur, Fundamentals of Molecular Evolution (Sunderland, Mass.: Sinauer, 1991), p. 69, Table 1.

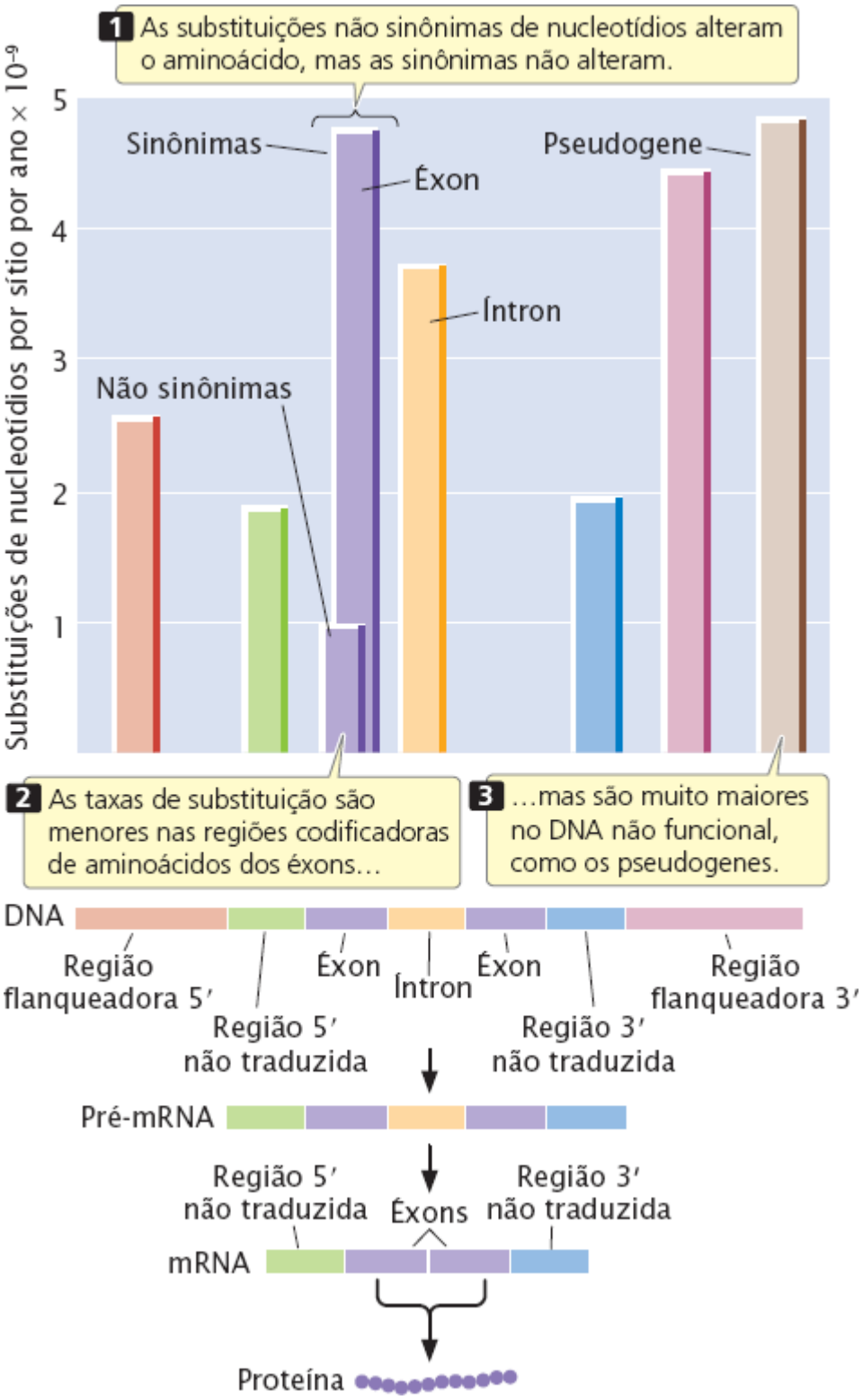


Figura 26.13 Diferentes partes dos genes evoluem em taxas diferentes. As taxas maiores de substituição de nucleotídios estão nas sequências que têm o menor efeito sobre a função da proteína.

As taxas de substituição são menores nas regiões 5′ e 3′ não traduzidas de um gene. Estas regiões são transcritas em RNA, mas não codificam aminoácidos. A região 5′ não traduzida tem um sítio de ligação ao ribossomo, que é essencial para tradução, e a região 3′ não traduzida tem sequências que podem atuar na estabilidade e na tradução do mRNA; então as substituições nestas regiões podem ter efeitos prejudiciais na aptidão (*fitness*) do organismo e podem não ser toleradas.

As menores taxas de substituição são observadas nas mudanças não sinônimas na região codificadora, pois estas substituições sempre alteram a sequência de aminoácidos da proteína e são prejudiciais. As altas taxas de substituições ocorrem nos pseudogenes, a maioria dos quais são cópias não funcionais duplicadas dos genes que adquiriram mutações. Tais genes não produzem mais um produto funcional, então as mutações nos pseudogenes têm pouco efeito na aptidão do organismo.

Em resumo, existe uma relação entre a função de uma sequência e sua taxa de evolução; taxas maiores são encontradas onde elas têm um efeito menor na função. Esta observação se encaixa com a hipótese de mutação neutra, que prevê que a variação molecular não é afetada pela seleção natural. Resolva o Problema 30

Relógio molecular

A hipótese de mutação neutra propõe que a mudança evolutiva em nível molecular ocorre principalmente pela fixação das mutações neutras pela deriva genética. A taxa na qual uma mutação neutra substitui outra depende apenas da taxa de mutação, a qual deve ser muito constante para qualquer gene específico. Se a taxa na qual uma proteína evolui é muito constante com o tempo, a mudança molecular que esta proteína sofreu pode ser usada como um **relógio molecular** para datar os eventos evolutivos.

Por exemplo, a enzima citocromo *c* poderia ser examinada em dois organismos conhecidos a partir de evidências fósseis por ter um ancestral comum de 400 milhões de anos. Ao determinar o número de diferenças nas sequências de aminoácidos do citocromo *c* em cada organismo, poderíamos calcular o número de substituições que teria ocorrido por sítio de aminoácido. A ocorrência de substituições dos 20 aminoácidos desde que os dois organismos divergiram indica uma taxa média de 5 substituições a cada 100 milhões de anos. Conhecer a velocidade do relógio molecular permite que usemos as mudanças moleculares no citocromo *c* para datar outros eventos evolutivos: se descobrirmos que o citocromo *c* nos dois organismos diferiu em substituições de 15 aminoácidos, nosso relógio molecular sugeriria que eles divergiram há 300 milhões de anos. Se considerarmos algum erro na nossa estimativa da taxa de substituição de aminoácidos, a análise estatística mostraria que o verdadeiro momento de divergência pode variar de 160 milhões a 440 milhões de anos. O relógio molecular foi proposto por Emile Zuckerkandl e Linus Pauling em 1965, como um possível meio para datar os eventos evolutivos com base nas moléculas nos organismos atuais. Vários estudos examinaram a taxa de mudança evolutiva nas proteínas (**Figura 26.14**) e nos genes e o relógio molecular foi muito usado para datar os eventos evolutivos quando não havia registro fóssil ou este era duvidoso. Por exemplo, os pesquisadores usaram um relógio molecular para estimar quando os tentilhões-de-darwin divergiram de um ancestral comum que originalmente colonizou as ilhas Galápagos. Esse relógio se baseia nas diferenças de sequência de DNA no gene do citocromo b. Eles concluíram que o ancestral dos tentilhões-de-darwin chegou nas ilhas Galápagos e começou a divergir há 2 a 3 milhões de anos. Os resultados de vários estudos mostraram que o relógio molecular nem sempre funciona em uma taxa constante, sobretudo em períodos mais curtos, e este método ainda é controverso.

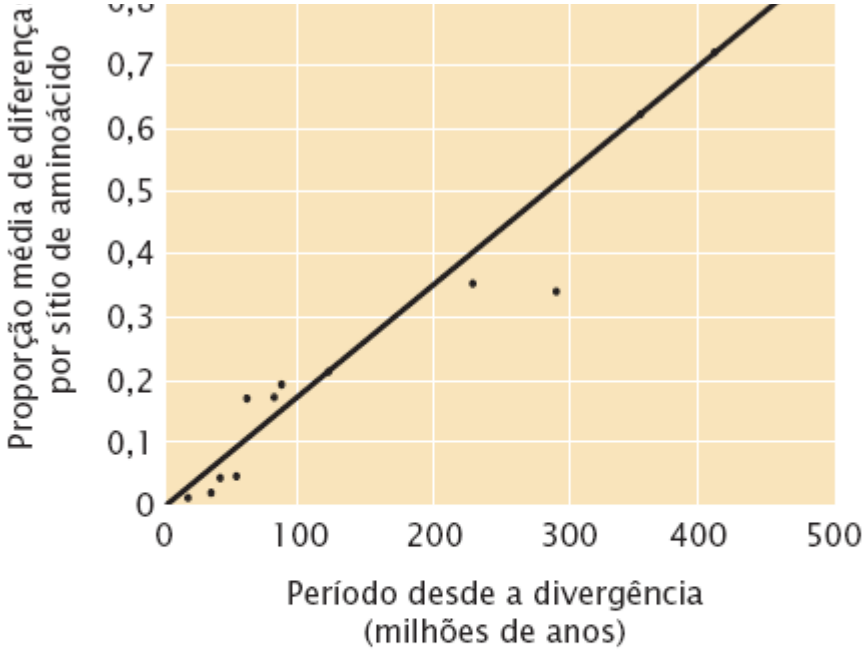
Conceitos

Diferentes genes e diferentes partes do mesmo gene evoluem em taxas diferentes. Estas partes dos genes que têm o menor efeito sobre a função tendem a evoluir nas taxas mais altas. A ideia do relógio molecular é que proteínas e genes individuais evoluem em uma taxa constante e que as diferenças nas sequências dos organismos atuais podem ser usadas para datar eventos evolutivos do passado.

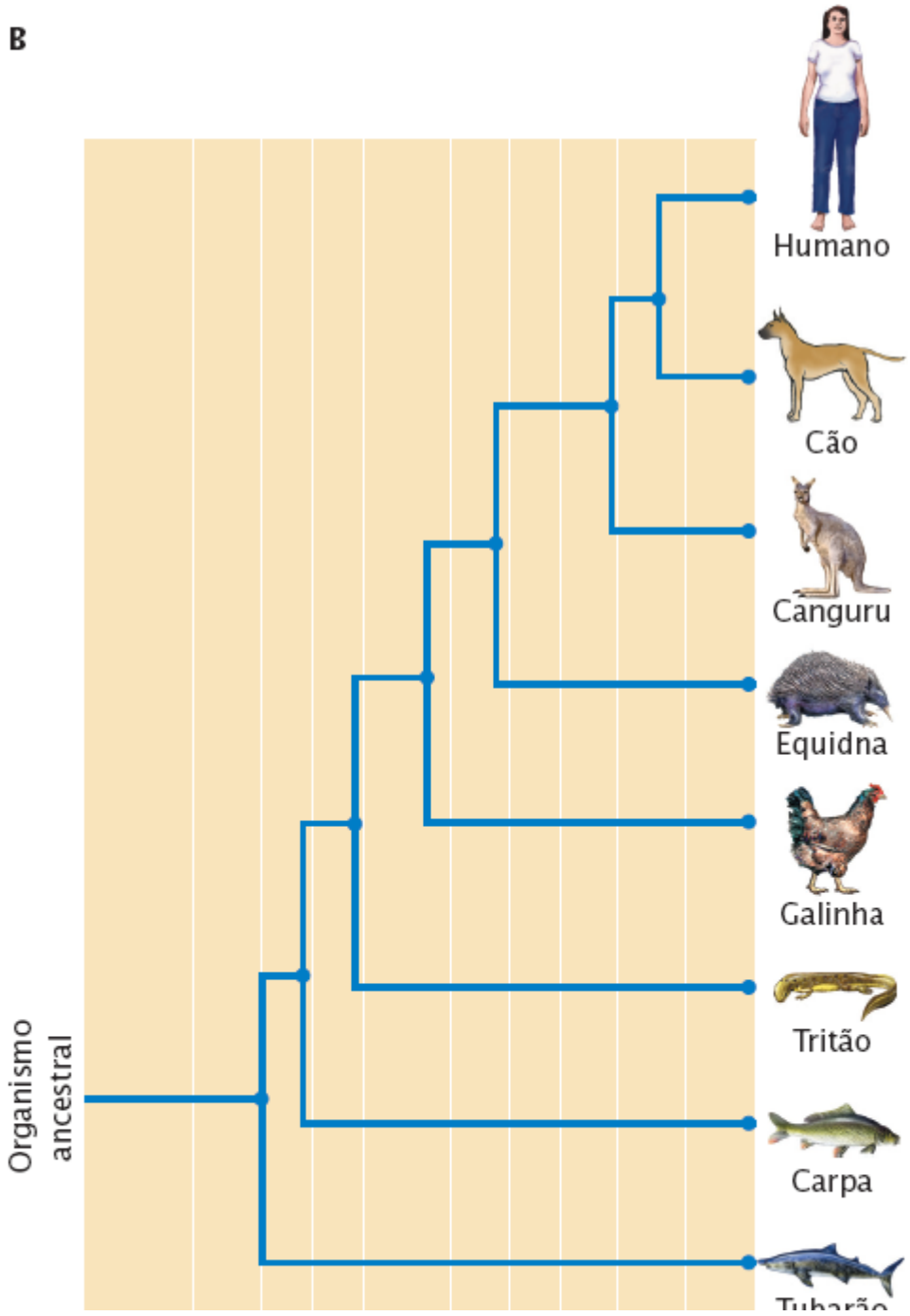
Checagem dos conceitos 7

Em geral, quais tipos de sequências devem exibir a mudança evolutiva?

a. Mudanças sinônimas nas regiões codificadoras de aminoácidos dos éxons.
b. Mudanças não sinônimas nas regiões codificadoras de aminoácidos dos éxons.
c. Íntrons.
d. Pseudogenes.


Proporção média de diferença por sítio de aminoácido
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0
100
200
300
400
500
Período desde a divergência (milhões de anos)

B


Humano
Cão
Canguru
Equidna
Galinha
Tritão
Carpa
Organismo ancestral

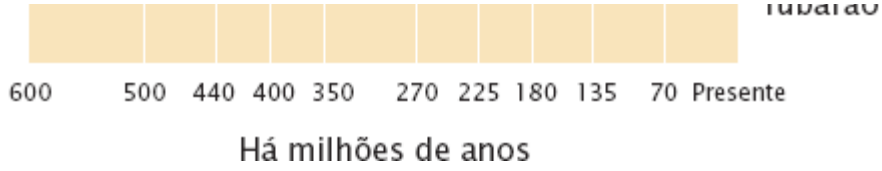


Figura 26.14 O relógio molecular baseia-se na suposição de uma taxa constante de mudança na proteína ou sequência de DNA. A. A relação entre a taxa de substituição de aminoácidos e o período desde a divergência baseia-se em parte nas sequências de aminoácidos da alfa-globina de oito espécies apresentadas na parte B. A taxa constante de evolução na proteína e sequências de DNA é usada como um relógio molecular para datar eventos evolutivos antigos. **B.** A filogenia de oito das espécies que são apresentados na parte A e seus períodos aproximados de divergência com base no registro fóssil.

Evolução por meio de mudanças na regulação gênica

Um dos desafios da biologia evolutiva é compreender a base genética da adaptação. Muitas mudanças evolutivas ocorrem com relativamente poucas diferenças genéticas. Por exemplo, os humanos e chimpanzés diferem muito na anatomia, fisiologia e comportamento e mesmo assim diferem em apenas 4% das suas sequências de DNA (ver a introdução do *Capítulo 17*). Os biólogos evolutivos assumem que muitas diferenças anatômicas não resultam a partir da evolução de novos genes, mas de diferenças relativamente pequenas de DNA que alteram a expressão dos genes existentes. Pesquisa recente na genética evolutiva se concentra em como a evolução ocorre pela alteração da expressão do gene.

Um exemplo de adaptação que ocorreu por mudanças nas sequências reguladoras é observado na evolução da pigmentação das moscas-da-fruta *Drosophila melanogaster* na África. A maioria das moscas-da-fruta têm uma cor canela-claro, mas algumas moscas em algumas populações da África exibem abdomes muito mais escuros. Estas moscas mais escuras ocorrem em regiões montanhosas em altitudes mais elevadas. De fato, 59% da variação de pigmentação entre as populações dentro da África Subsaariana pode ser explicada pelas diferenças na elevação (**Figura 26.15**). Os pesquisadores demonstraram que estas diferenças são geneticamente determinadas e que a seleção natural favoreceu a pigmentação mais escura em altitudes mais elevadas. As populações de altitudes mais elevadas são expostas a menores temperaturas e supõe-se que a pigmentação mais escura ajude as moscas a absorver mais a radiação solar e regular melhor sua temperatura do corpo nesses ambientes.

Como as moscas das altitudes elevadas desenvolveram uma cor mais escura? Estudos genéticos indicam que a pigmentação abdominal escura observada nas moscas a partir destas populações resulta da variação dentro ou próximo de um *locus* chamado *ebony*. O *locus ebony* codifica uma enzima multifuncional que produz um exoesqueleto de cor amarela; sua ausência produz um fenótipo escuro. O sequenciamento do *locus ebony* das moscas de populações de cores clara e escura não encontrou diferenças na região codificadora do gene *ebony*. Entretanto, a análise molecular relevou uma marcante redução do mRNA *ebony* nas moscas mais escuras, sugerindo de que as diferenças na pigmentação não são devido a mutações no gene *ebony* propriamente dito, mas na sua expressão. Outra investigação detectou diferenças genéticas dentro de um acentuador que está cerca de 3.600 *upstream* do gene *ebony*. As moscas escuras e claras diferem em cerca de 120 nucleotídios dispersos em mais de 2.400 pb do acentuador. Entretanto, ao criar experimentalmente acentuadores com diferentes combinações destes mutantes, os pesquisadores determinaram que cinco das mutações são responsáveis pela maioria das diferenças na pigmentação.

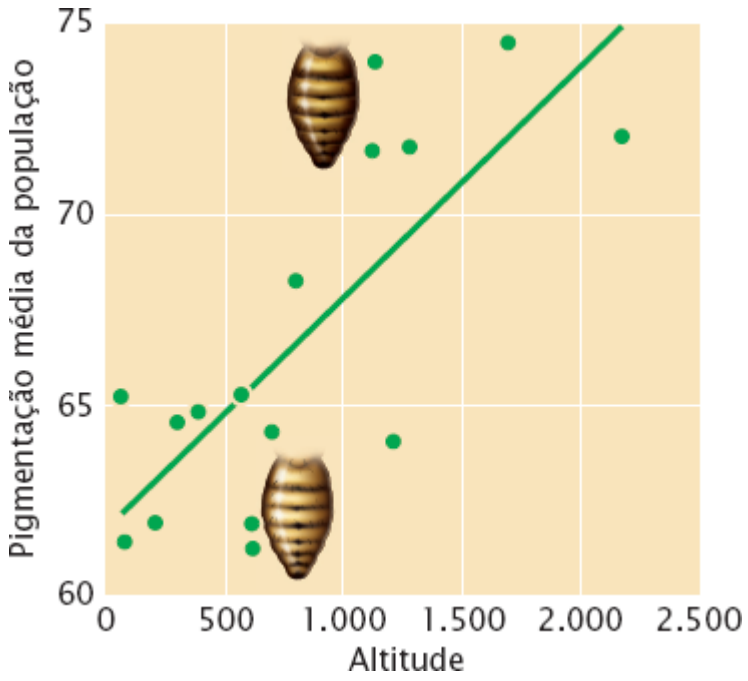


Figura 26.15 As populações africanas subsaarianas de mosca-da-fruta *Drosophila melanogaster* exibem uma associação positiva entre a pigmentação e altitude.

Esses estudos sugeriram que as populações de altitudes elevadas acumularam múltiplas mutações no acentuador, o que reduziu a expressão do *locus ebony* e causou a pigmentação mais escura. Outra análise sugeriu que estas mutações eram adicionadas em sequência. Algumas destas mutações estão espalhadas por toda a África; supõe-se que estas mutações existentes foram favorecidas por seleção natural nas populações de altas altitudes e aumentaram em frequência porque ajudaram as moscas a termorregular nos ambientes mais frios. Outras mutações são observadas apenas em populações de altitudes elevadas, sugerindo que provavelmente surgiram como novas mutações dentro destas populações e foram rapidamente favorecidas pela seleção natural.

Evolução do genoma

O rápido crescimento dos dados de sequência disponíveis no banco de dados de DNA é uma fonte de discernimento nos processos evolutivos. As sequências de genoma inteiro também estão fornecendo novas informações sobre como os genomas evoluem e os processos que moldam o tamanho, a complexidade e a organização dos genomas.

Embaralhamento de éxons. Muitas proteínas são compostas por grupos de aminoácidos, chamados domínios, que especificam funções discretas ou contribuem para a estrutura molecular de uma proteína. Por exemplo, no Capítulo 16, consideramos os domínios de ligação do DNA das proteínas que regulam a expressão gênica. As análises de sequências de genes a partir de organismos eucarióticos indicam que os éxons codificam domínios funcionais discretos de proteínas.

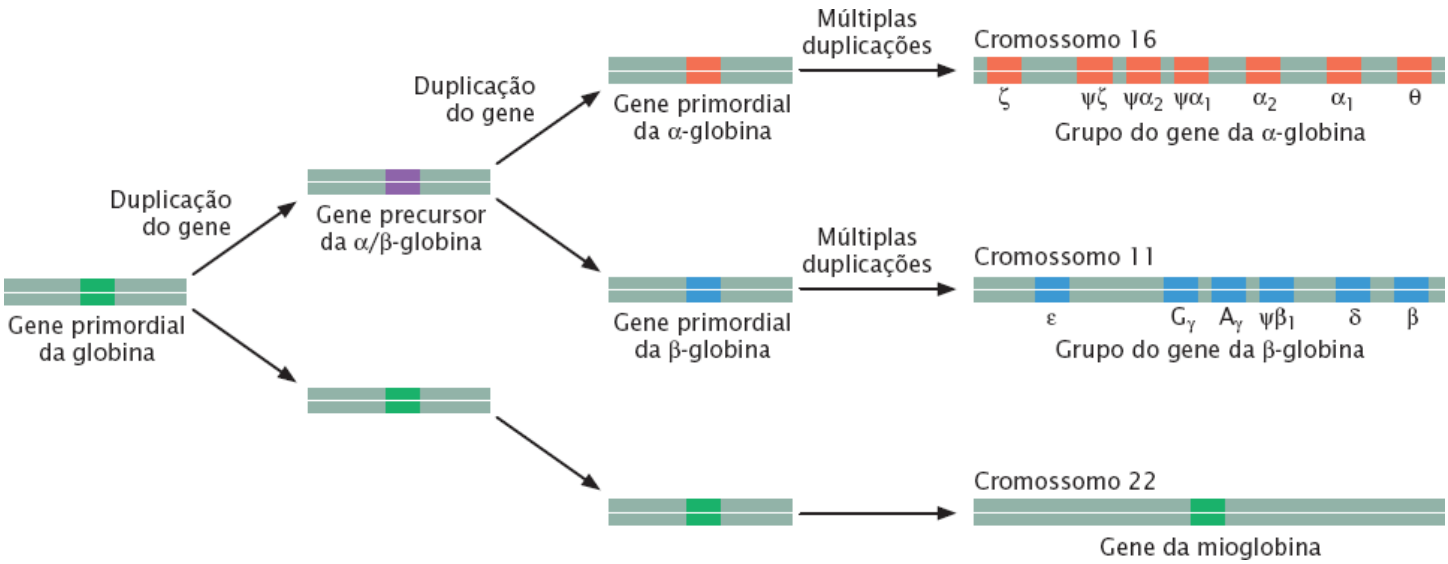


Figura 26.16 Os genes da globina humana constituem uma família de multigenes que evoluiu a partir de sucessivas duplicações de genes.

Alguns genes se alongaram e desenvolveram novas funções quando um ou mais éxons duplicaram ou sofreram divergência. Por exemplo, o gene da albumina sérica humana é composto por três cópias de uma sequência que codifica um domínio de proteína de 195 aminoácidos. Além disso, os genes que codificam as imunoglobulinas humanas sofreram duplicações em *tandem* repetidas, criando muitos segmentos *V*, *J*, *D* e *C* semelhantes (ver *Capítulo 22)*, que permitem que o sistema imunológico responda a quase todas substâncias estranhas que entram no corpo.

Uma comparação das sequências de DNA de diferentes genes revela que novos genes evoluem repetidamente por um processo chamado de **embaralhamento de éxons**, no qual os éxons de diferentes genes são trocados, criando genes que são mosaicos de outros genes. Por exemplo, o ativador do plasminogênio de tecido (TPA) é uma enzima com quatro domínios de três tipos diferentes, chamados de *kringle*, fator de crescimento e *finger*. Cada domínio é codificado por um éxon diferente. O gene para TPA adquiriu seus éxons a partir de outros genes que codificam diferentes proteínas: o éxon *kringle* veio do gene do plasminogênio, o éxon do fator de crescimento veio do gene do fator de crescimento epidérmico, e o éxon *finger* veio do gene da fibronectina. O mecanismo pelo qual o embaralhamento de éxons ocorre é pouco compreendido, mas novas proteínas com diferentes combinações de funções codificadas por outros genes aparentemente evoluíram de forma repetida por este mecanismo.

Duplicação do gene. Novos genes também evoluíram pela duplicação de genes inteiros e sua subsequente divergência. Este processo cria **famílias de multigenes**, conjuntos de genes com sequências semelhantes, mas que codificam produtos diferentes.

Por exemplo, os humanos têm 13 genes diferentes encontrados nos cromossomos 11 e 16 que codificam moléculas semelhantes a globina, que participam no transporte de oxigênio (**Figura 26.16**). Todos estes genes têm uma estrutura semelhante, com três éxons separados por dois íntrons e supõe-se que evoluíram pela duplicação repetida e divergência de um único gene de globina em um ancestral distante. Acredita-se que este gene ancestral seja mais semelhante ao gene da mioglobina atual e duplicado pela primeira vez para produzir um gene precursor de α/β-globina e o gene da mioglobina. O gene da α/β-globina sofreu outra duplicação para dar origem a um gene primordial da α-globina e um gene primordial da β-globina. Duplicações subsequentes levaram a múltiplos genes da α-globina e β-globina. Da mesma forma, os vertebrados têm quatro grupos de genes *Hox*, cada um com 9 a 11 genes. Os genes *Hox* são importantes no desenvolvimento (ver *Capítulo 22*).

Algumas famílias gênicas incluem genes que foram organizados em *tandem* no mesmo cromossomo; outros estão dispersos entre diferentes cromossomos. A duplicação de gene é uma ocorrência comum nos genomas eucarióticos, por exemplo, cerca de 5% do genoma humano é composto por segmentos duplicados.

A duplicação de genes fornece um mecanismo para a adição de novos genes com novas funções; após um gene duplicar, existem duas cópias da sequência, uma das quais está livre para mudar e potencialmente assumir uma nova função. A cópia extra do gene pode, por exemplo, ser ativada em um momento diferente no desenvolvimento ou ser expressa em um diferente tecido ou até divergir e codificar uma proteína com diferentes aminoácidos. Entretanto, o destino mais comum da duplicação de genes é que uma cópia adquire uma mutação que o torna não funcional, dando origem a um pseudogene. Os pseudogenes são comuns nos genomas de eucariotos complexos; estima-se que genoma humano tenha mais de 20.000 pseudogenes.

Duplicação do genoma inteiro. Além da duplicação de genes individuais, os genomas inteiros de alguns organismos duplicaram no passado. Por exemplo, uma comparação do genoma da levedura *Saccharomyces cerevisiae* com os genomas de outros fungos relevou que a *S. cerevisiae* ou um dos seus ancestrais imediatos sofreu uma duplicação de genoma inteiro, gerando duas cópias de cada gene. Muitas das cópias subsequentemente adquiriram novas funções, outras adquiriram mutações que destruíram a função original e então divergiram em sequências de DNA aleatórias. A duplicação do genoma inteiro pode ocorrer pela poliploidia.

Durante sua evolução, as plantas sofreram várias duplicações de genoma inteiro pela poliploidia. Enquanto a poliploidia é menos comum nos animais, evidências genéticas sugerem que ocorreram vários eventos de duplicação de genoma inteiro durante a evolução animal. Em 1970, Susumu Ohno propôs que os primeiros vertebrados sofreram dois ciclos de duplicação de genoma. Chamada de hipótese 2R, esta ideia é controversa, mas dados recentes de sequenciamento de genoma forneceram provas para a hipótese.

Transferência horizontal de genes. Tradicionalmente, os cientistas consideram que os organismos adquirem seus genomas pela transmissão vertical, transferência pela reprodução das informações genéticas dos genitores para os descendentes e a maioria das árvores filogenéticas considera a transmissão vertical da informação genética. Achados de estudos de sequência de DNA revelam que as sequências de DNA são trocadas em alguns casos por uma transferência horizontal de genes, na qual o DNA é transferido entre indivíduos de diferentes espécies (ver *Capítulo 9*). Esse processo é especialmente comum entre as bactérias e existem vários casos documentados nos quais os genes são transferidos de bactérias para eucariotos. A extensão da transferência horizontal de genes entre organismos eucarióticos é controversa, com poucos casos bem documentados. A transferência horizontal de genes pode encobrir relações filogenéticas e dificultar a reconstrução das árvores filogenéticas.

Um caso aparente de transferência horizontal de genes entre eucariotos é a presença em alguns afídeos de genes para enzimas que sintetizam carotenoides. Os carotenoides são compostos coloridos produzidos por bactérias, archaea, fungos e plantas. Muitos animais também têm carotenoides, mas não têm as enzimas necessárias para produzir os compostos; em quase todos os casos os animais obtêm os carotenoides da alimentação.

Os afídeos, pequenos insetos que se alimentam de plantas, têm carotenoides, que são responsáveis pelas diferenças de cor entre e dentro da espécie. Alguns afídeos são verdes e têm α-, β- e γ-caroteno, que são todos carotenoides amarelos. Outros afídeos são vermelhos ou marrons e têm licopeno ou toruleno, carotenoides vermelhos. Uma espécie, o afídeo da ervilha (*Acyrthosiphon pisum*), tem indivíduos verdes e vermelhos e estas diferenças são geneticamente herdadas. Muitos pesquisadores consideravam que as diferenças de cor fossem causadas pelos carotenoides obtidos pela alimentação dos afídeos.

Recentemente, os pesquisadores sequenciaram o genoma inteiro do *A. pisum*, que forneceu a oportunidade de determinar se os afídeos têm seus próprios genes para a síntese de carotenoide. O exame das sequências genômicas revelou que os afídeos da ervilha tinham vários genes que codificam enzimas sintetizadoras de carotenoide. E interessante, estes genes estão intimamente relacionados com os genes sintetizadores de carotenoide encontrados em alguns fungos. Evidências sugerem que no passado distante, um afídeo adquiriu seus genes para carotenoide de um fungo pela transferência horizontal de genes e então transmitiu os genes para outros afídeos pela transmissão vertical.

Conceitos

Novos genes podem evoluir pela duplicação de éxons, embaralhamento de éxons, duplicação de genes e duplicação de genomas inteiros. Os genes podem ser transmitidos entre organismos distantes pela transferência horizontal de genes.

Resumo dos conceitos

- A evolução é a mudança genética que ocorre em um grupo de organismos. É um processo de duas etapas: (1) surge a variação genética e (2) ocorre mudança na frequência das variantes genéticas
- A anagênese é mudança em uma única linhagem, a cladogênese é a divisão de uma linhagem em duas
- Os métodos moleculares oferecem várias vantagens para o estudo da evolução
- O uso de eletroforese de proteínas para estudar variação genética nas populações naturais mostrou que a maioria das populações naturais tem substancial variação genética nas suas proteínas. A hipótese de mutação neutra propõe que a variação molecular é seletivamente neutra e é moldada em grande parte por mutação e deriva genética. A hipótese de equilíbrio propõe que a variação molecular seja mantida em grande parte pela seleção de equilíbrio
- A variação na sequência de DNA pode ser analisada por meio de polimorfismos de comprimento do fragmento de restrição, microssatélites e dados de sequenciamento direto
- Uma espécie pode ser definida como um grupo de organismos capazes de cruzar um com outro e estão isolados do ponto de vista reprodutivo dos mesmos de outra espécie
- As espécies são impedidas de trocar genes pelo isolamento reprodutivo pré-zigótico ou pós-zigótico
- A especiação alopátrica surge quando uma barreira geográfica evita o fluxo gênico entre duas populações. Com o passar do tempo, as duas populações adquirem diferenças genéticas que podem levar a mecanismos de isolamento reprodutivo
- A especiação simpátrica surge quando o isolamento reprodutivo existe na ausência de barreira geográfica. Ela pode surgir sob circunstâncias especiais
- Algumas espécies surgem apenas após as populações sofrerem consideráveis diferenças genéticas, outras surgem após ocorrerem mudanças em apenas alguns genes
- As relações evolutivas (uma filogenia) podem ser representadas por uma árvore filogenética, composta por nós que representam organismos e ramos que representam suas conexões evolutivas
- As abordagens para construir as árvores filogenéticas incluem a abordagem por distância, a abordagem da máxima parcimônia e a abordagem da máxima verossimilhança e métodos bayesianos
- Diferentes pares do genoma mostram diferentes quantidades de variação genética. Em geral, estas partes que têm o menor efeito sobre a função tendem a evoluir em taxas maiores
- A hipótese de relógio molecular propõe uma taxa constante de substituição de nucleotídios, fornecendo um meio para datar os eventos evolutivos ao observar as diferenças de nucleotídios entre os organismos
- A evolução do genoma ocorre graças a duplicação e embaralhamento de éxons, duplicação dos genes para formar famílias de genes, duplicação do genoma inteiro e transferência horizontal de genes entre organismo.

Termos importantes

Anagênese
Árvore filogenética
Árvore filogenética com raiz
Árvore gênica
Cladogênese
Colapso do híbrido
Conceito biológico de espécie
Embaralhamento de éxons
Especiação
Especiação alopátrica
Especiação simpátrica
Espécie
Esterilidade híbrida
Evolução
Famílias de multigenes
Filogenia

Hipótese de equilíbrio
Hipótese de mutação neutra
Inviabilidade híbrida
Isolamento comportamental
Isolamento ecológico
Isolamento gamético
Isolamento mecânico
Isolamento temporal
Mecanismo de isolamento reprodutivo
Mecanismo de isolamento reprodutivo pós-zigótico
Mecanismo de isolamento reprodutivo pré-zigótico
Nó (ponto de ramificação)
Ramo
Relógio molecular

Respostas da Checagem dos conceitos

1. Primeiro, surge a variação genética. Então várias forças evolutivas provocam mudanças na frequência das variantes genéticas.
2. c.
3. Os microssatélites são muito variáveis entre os indivíduos. Eles podem ser ampliados com o uso de PCR, e podem ser detectados com uma pequena quantidade de DNA inicial. Finalmente, a detecção e a análise dos microssatélites podem ser automatizadas.
4. b.
5. A deriva genética pode desencadear alterações nas frequências alélicas das populações e levar a diferenças genéticas entre as populações. A diferenciação genética é a causa do isolamento reprodutivo pós-zigótico e pré-zigótico entre populações que leva à especiação.
6. c.
7. b.

Questões de compreensão

Seção 26.1

1. Como é definida a evolução biológica?
2. Quais são as duas etapas no processo de evolução?
3. Qual a diferença entre anagênese e cladogênese?

Seção 26.2

4. Por que a variação da proteína, como revelada pela eletroforese, subestima a variação genética verdadeira?
5. Quais são algumas das vantagens em usar os dados moleculares nos estudos evolutivos?
6. Qual é a diferença principal entre a hipótese de mutação neutra e a hipótese de equilíbrio?
7. Descreva alguns dos métodos usados para estudar a variação no DNA.

Seção 26.3

8. O que é o conceito biológico de espécies?
9. Qual é a diferença entre os mecanismos de isolamento reprodutivo pré-zigótico e pós-zigótico? Liste diferentes tipos de cada mecanismo.
10. Qual é a diferença básica entre os modos de especiação alopátrica e simpátrica?
11. Descreva resumidamente o processo de especiação alopátrica.
12. Quais são algumas das dificuldades com a especiação simpátrica?
13. Explique resumidamente como a troca de frutas do espinheiro para maçãs levou a diferenciação genética e isolamento reprodutivo parcial da mosca *Rhagoletis pomonella*.

Seção 26.4

14. Desenhe uma árvore filogenética simples e identifique um nó, um ramo e um grupo externo.
15. Descreva resumidamente as diferenças entre abordagem por distância, a abordagem de máxima parcimônia e a abordagem de máxima verossimilhança para a reconstrução das árvores filogenéticas.

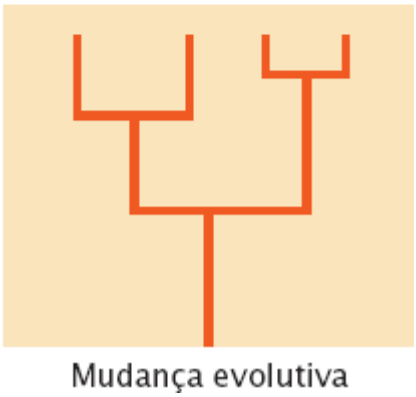
Seção 26.5

16. Destaque as diferentes taxas de evolução que são observadas em diferentes partes de um gene codificador de proteína. O que pode explicar estas diferenças?
17. O que é relógio molecular?
18. O que é embaralhamento do éxon? Como pode levar à evolução de novos genes?
19. O que é uma família multigênica? Quais processos produzem as famílias multigênicas?
20. Defina a transferência horizontal de genes. Quais problemas esta transferência gera para os biólogos evolutivos?

Questões e problemas aplicados

Seção 26.1

*21. As ilustrações a seguir representam dois padrões diferentes de evolução. Discuta resumidamente as diferenças nestes dois padrões, sobretudo com relação ao papel da cladogênese na mudança evolutiva.



Seção 26.3

*22. Quais dos mecanismos de isolamento listados no Quadro 26.2 evoluíram parcialmente entre as raças hospedeiras maçã e espinheiro da mosca-da-fruta *Rhagoletis pomonella*?

ANÁLISE DE DADOS

23. Analisamos a evolução simpátrica dos mecanismos de isolamento reprodutivo nas raças hospedeiras da mosca-da-fruta *Rhagoletis pomonella* (também chamada mosca-da-maçã). A vespa *Diachasma alloeum* parasita essas moscas, depositando seus ovos nas larvas das moscas. As vespas imaturas saem dos ovos e se alimentam das larvas da mosca. A pesquisa realizada por Andrew Forbes e colaboradores (Forbes *et al*. 2009. *Science* 323:776-779) demonstrou que as vespas que parasitam *R. pomonella* são geneticamente diferentes das que parasitam *R. pomonella* do espinheiro (pilriteiro). Eles também descobriram que as vespas que atacam as moscas-da-maçã são atraídas pelo odor das maçãs, enquanto as vespas que atacam a mosca do espinheiro são atraídas pelo odor do fruto do espinheiro. Proponha uma

hipótese para explicar como as diferenças genéticas podem ter evoluído entre as vespas que parasitam as duas raças de *R. pomonella*. Como estas diferenças podem ter levado à especiação nas vespas?

24. Qual das espécies na Figura 26.9 é um poliploide?
    a. *Spartina maritima*.
    b. *Spartina alterniflora.*
    c. Híbrido entre *Spartina maritima* e *Spartina alterniflora*.
    d. *Spartina anglica*.
    e. Opções c e d.

Seção 26.4

25. Quantas árvores com raiz são teoricamente possíveis para um grupo de sete organismos? E quantas para 12 organismos?
26. Alinhe as sequências a seguir para maximizar sua similaridade. Qual o número mínimo de etapas evolutivas para separar estas duas sequências?

TTGCAAAC
TGAAACTG

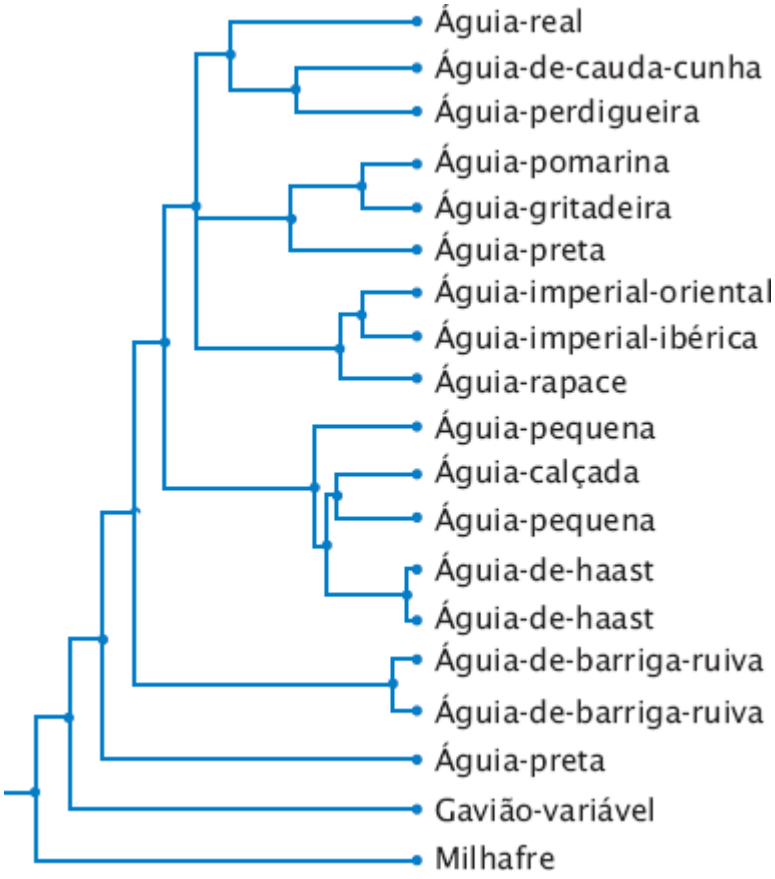
*27. Micheal Bunce e seus colegas na Inglaterra, Canadá e EUA extraíram e sequenciaram o DNA mitocondrial de fósseis da águia-de-haast, uma águia gigante que foi extinta há 700 anos quando os humanos chegaram pela primeira vez na Nova Zelândia (M. Bunce *et al*. 2005. *PLOS Biology* 3:44-46). Usando as sequências de DNA mitocondrial obtidas de águias vivas e dos fósseis da águia-de-haast, eles criaram a árvore filogenética a seguir. Nesta árvore filogenética, identifique: (a) todos os nós terminais; (b) todos os nós internos; (c) um exemplo de um ramo e (d) o grupo externo.

28. Com base na filogenia dos tentilhões-de-darwin apresentada na Figura 26.7, indique quais duas espécies de cada um dos grupos a seguir terá maior similaridade genética.
    a. *Camarhynchus parvulus, Camarhynchus psittacula, Camarhynchus pallida.*
    b. *Camarhynchus parvulus, Camarhynchus pallida, Platyspiza crassirostris.*
    c. *Geospiza difficilis, Geospiza conirostris, Geospiza scandens.*
    d. *Camarhynchus parvulus, Certhidea fusca, Pinaroloxias inornata.*

Seção 26.5

29. Considere que um dos genes apresentados no Quadro 26.4 tinha taxas de substituições não sinônimas e sinônimas. O que isto pode sugerir sobre a evolução deste gene?

*30. Com base nas informações fornecidas na Figura 26.13, os íntrons ou regiões 3′ não traduzidas de um gene têm as taxas mais elevadas de substituição de nucleotídio? Explique o motivo.

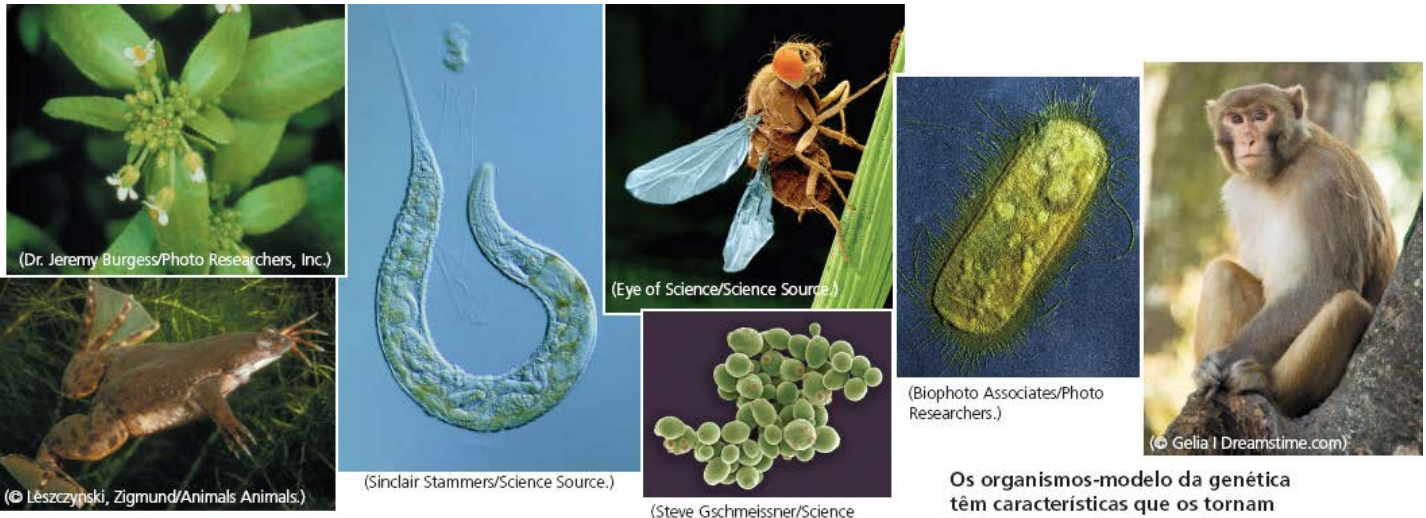


(De M. Bunce et al., PLOS Biology 3:44-46, 2005.)

Questões desafiadoras

Seção 26.3

31. Explique por que a seleção natural promove a evolução de mecanismos de isolamento reprodutivo pré-zigótico se já existirem mecanismos de isolamento reprodutivo pós-zigótico, mas nunca consegue provocar a evolução dos mecanismos de isolamento reprodutivo pós-zigótico.
32. A poliploidia é muito comum nas plantas florescentes: aproximadamente 40% de todas as espécies de plantas florescentes são poliploides. Embora a poliploidia exista em muitos grupos animais diferentes, ela é muito menos comum. Por que a poliploidia é mais comum nas plantas do que nos animais? Indique um ou mais possíveis motivos.



Os organismos-modelo da genética têm características que os tornam úteis para os estudos genéticos. Aqui estão apresentados os organismos comumente usados nos estudos genéticos.

Guia de Referência para Organismos-Modelo da Genética

O que Lou Gehrig, o melhor jogador de beisebol na história da liga nacional, e Stephen Hawking, o mais famoso físico teórico do mundo, têm em comum? Ambos sofrem ou sofreram de esclerose lateral amiotrófica (ELA, também conhecida como doença de Lou Gehrig), uma doença neurológica degenerativa que leva à fraqueza progressiva e atrofia dos músculos esqueléticos. Vale a pena mencionar que a pesquisa recente revela que, na verdade, alguns atletas com sintomas de ELA sofreram traumatismo cranioencefálico (TCE) durante a prática desportiva; então Lou Gehrig não poderia ser homenageado com o nome da doença.

A maioria dos casos de ELA é esporádica, aparecem em pessoa sem história familiar da doença. Entretanto, cerca de 10% dos casos têm caráter familiar e são herdados. Em 2004, os geneticistas descobriram uma grande família no Brasil com múltiplos casos de ELA. A análise genética revelou que a esclerose lateral amiotrófica nesta família é causada por uma mutação em um gene chamado *VABP*, que codifica uma proteína de membrana associada com as vesículas.

Para melhor compreender como as mutações no *VABP* levam às manifestações da ELA, os geneticistas se voltaram para uma cobaia improvável, a mosca-da-fruta *Drosophila melanogaster*. As moscas-da-fruta não têm ELA, mas têm um gene muito semelhante ao *VABP*. Usando uma ampla gama de técnicas que foram desenvolvidas para a manipulação genética de moscas-da-fruta, os geneticistas criaram moscas transgênicas com a sequência mutante do gene *VABP* que provoca a síndrome ELA nos seres humanos. Estas moscas são um modelo de doença para ELA e estão sendo usadas para melhor compreender o que o gene *VABP* normalmente faz e como sua alteração pode levar a ELA.

O campo da genética tem sido influenciado e moldado por alguns organismos-chave, chamados de organismos-modelo da genética, cujas características os tornam particularmente responsivos aos estudos genéticos. O uso da *Drosophila* para estudar a ELA nos humanos ilustra o valor desta abordagem. Como as características dos sistemas genéticos são comuns a muitos organismos, a pesquisa realizada em uma espécie pode ser uma fonte de informação sobre os sistemas genéticos de outras espécies. Esta associação de função genética significa que os geneticistas podem concentrar seus esforços em organismos-modelo que são fáceis de trabalhar e podem gerar resultados.

Os organismos-modelo da genética têm ciclos de vida e características genômicas que os tornam bem adequados para o estudo e a análise genética. Algumas características-chave dos muitos organismos-modelo da genética incluem:

- Um curto tempo de geração e, assim, várias gerações de cruzamentos genéticos podem ser examinadas em um período de tempo razoável
- A produção de numerosos descendentes, o que facilita observar as razões genéticas
- A capacidade de fazer e controlar os cruzamentos genéticos no organismo
- A capacidade de ser cultivado em um ambiente laboratorial, exigindo pouco espaço e poucos recursos para manter

- A disponibilidade de numerosas variantes genéticas
- Um conjunto de conhecimentos acumulados sobre os seus sistemas genéticos.

Recentemente, os genomas de muitos organismos-modelo da genética foram completamente sequenciados, facilitando muito seu uso na pesquisa genética. Nem todos os organismos-modelo têm todas estas características.

Entretanto, cada organismo-modelo da genética tem uma ou mais características que o tornam útil para a análise genética. Por exemplo, o milho não pode ser cultivado com facilidade no laboratório (e, em geral, não é) e tem um tempo de geração relativamente longo, mas ele produz numerosos descendentes e existem muitas variantes genéticas do milho disponíveis para estudo.

Este guia de referência destaca seis organismos-modelo da genética com papéis importantes no desenvolvimento da genética: mosca-da-fruta (*Drosophila melanogaster*), bactéria (*Escherichia coli*), nematelminto (*Caenorhabditis elegans*), planta *Arabidopsis thaliana*, camundongo doméstico (*Mus musculus*) e levedura (*Saccharomyces cerevisiae*). Esses seis organismos são amplamente usados em pesquisa e ensino da genética. Vários outros organismos também são usados como sistemas modelos na genética, incluindo milho (*Zea mays*), peixe-zebra (*Danio rerio*), rã-de-unhas-africana (*Xenopus lavis*), bolor de pão (*Neurospora crassa*), rato (*Rattus norvegicus*) e macaco *rhesus* (*Macaca mulatta*), só para citar alguns.

Mosca-da-fruta *Drosophila melanogaster*

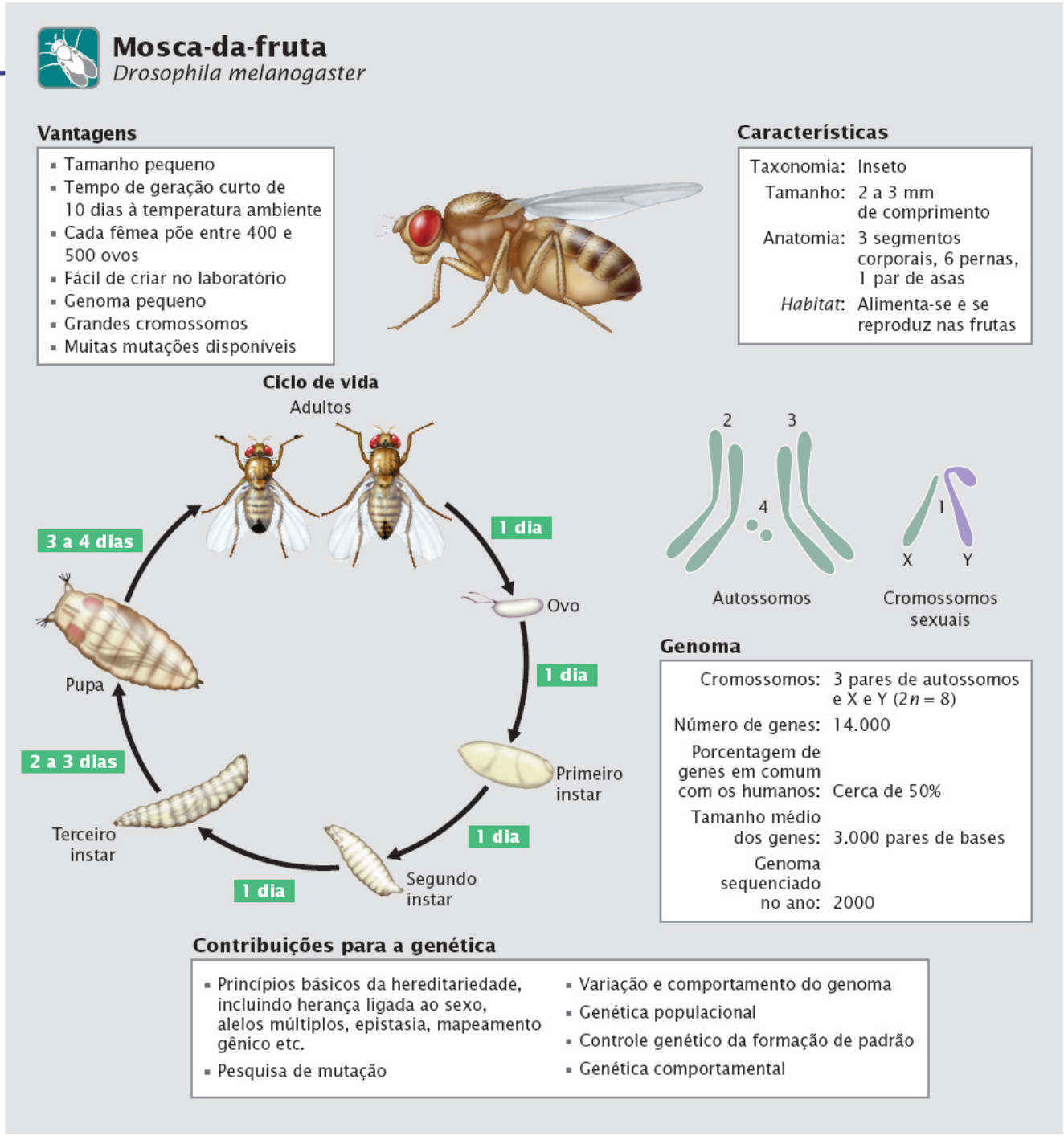
A *Drosophila melanogaster*, uma mosca-da-fruta, foi um dos primeiros organismos usados para análise genética e é um dos mais usados e conhecidos do ponto de vista genético. Sua participação foi importante nos estudos de ligação, epistasia, genética do cromossomo, desenvolvimento, comportamento e evolução. Como todos os organismos usam um sistema genético comum, a compreensão de tal processo, como a replicação ou a transcrição nas moscas-da-fruta, ajuda-nos a compreender estes mesmos processos nos humanos e outros eucariotos.

Drosophila é um gênero de mais de 1.000 espécies descritas de pequenas moscas (cerca de 1 a 2 mm de comprimento) que se alimentam e reproduzem em frutos, embora raramente causem dano e não sejam consideradas pragas do ponto de vista econômico. A mosca-da-fruta mais conhecida e amplamente estudada é a *D. melanogaster*, mas os estudos genéticos se estendem a muitas outras espécies do gênero também. *D. melanogaster* foi empregada pela primeira vez nos laboratórios de biologia em torno de 1900. Após fazer os primeiros experimentos de reprodução com camundongos e ratos, Thomas Hurt Morgan começou a usar as moscas-da-fruta em estudos experimentais de hereditariedade na Columbia University. O laboratório de Morgan, localizado no último andar do Schermerhorn Hall, se tornou conhecido como o Quarto da Mosca (ver Figura 4.9 B). Dizer que o Quarto da Mosca era inexpressivo é pouco. Na sala apertada, com apenas 4,8 × 7 metros, havia oito mesas, cada uma ocupada por um estudante e seus experimentos. Os primitivos equipamentos de laboratório eram pouco mais do que frascos de leite para criar as moscas e lentes portáteis para observar seus traços. Posteriormente, os microscópios substituíram as lentes portáteis e foram incluídas incubadoras primitivas para manter as culturas de moscas, mas até estes acréscimos tiveram pouco efeito para aumentar a sofisticação física do laboratório. Morgan e seus estudantes não se protegiam: as baratas eram abundantes (vivendo sobre a comida derramada da *Drosophila*), frascos sujos de leite enchiam a pia, bananas maduras – os alimentos para as moscas – eram penduradas no teto, e moscas-da-fruta que escapavam pairavam em todos os lugares. Apesar das suas limitações físicas, o Quarto da Mosca foi a fonte de parte de algumas das pesquisas mais importantes na história da biologia. Existia um entusiasmo diário entre os estudantes, alguns dos quais chegavam ao laboratório sem estarem formados. As dimensões reduzidas do local facilitavam a informalidade e o fluxo livre de ideias. Morgan e o Quarto da Mosca ilustram a grande importância da "atmosfera" para produzir ciência de qualidade. Morgan e seus estudantes acabaram elucidando muitos princípios básicos da hereditariedade, incluindo herança ligada ao sexo, epistasia, alelos múltiplos e mapeamento gênico, graças a *Drosophila*.

Vantagens de *D. melanogaster* como organismo-modelo da genética. O uso disseminado da *Drosophila* nos estudos de genética não é por acaso. A mosca-da-fruta tem várias características que a tornam ideal para investigações genéticas. Comparada com outros organismos, tem um tempo de geração relativamente curto; as moscas-da-fruta completam uma geração inteira em 10 dias à temperatura ambiente; assim, várias gerações podem ser estudadas em algumas semanas. Embora a *D. melanogaster* tenha um tempo de geração curto, ela tem um ciclo de vida complexo, passando por vários estágios de desenvolvimento diferentes, incluindo ovo, larva, pupa e adulto. Uma mosca-da-fruta fêmea é capaz de cruzar 8 horas após surgir e, em geral, começa a pôr ovos após 2 dias. As moscas-da-fruta também produzem muitos descendentes, colocando entre 400 e 500 ovos em um período de 10 dias. Dessa forma, são obtidos muitos descendentes a partir de um único cruzamento genético.

Outra vantagem é a facilidade para criar estas moscas em laboratório. Em geral elas são criadas em pequenos frascos ou garrafas de vidro e são alimentadas por ração pastosa, de fácil preparo, composta por bananas, farinha de milho e melado. Os machos e fêmeas são diferenciados sem dificuldade e as fêmeas virgens são facilmente isoladas, facilitando os cruzamentos genéticos. As moscas são pequenas, exigindo pouco espaço – muitas centenas podem ser criadas em uma garrafa de 250 mℓ –, mas são grandes o suficiente para muitas mutações serem observadas com facilidade com o uso de lente ou um microscópio de dissecação.

Finalmente, *D. melanogaster* é o organismo preferido para muitos geneticistas porque tem um genoma relativamente pequeno com 175 milhões de pares de bases (pb) de DNA, que é apenas cerca de 5% do tamanho do genoma humano. Ela tem quatro pares de cromossomos, três pares de autossomos e um par de cromossomos sexuais. O cromossomo X (designado como cromossomo 1) é grande e acrocêntrico, enquanto o cromossomo Y é grande e submetacêntrico, embora tenha muito pouca informação genética. Os cromossomos 2 e 3 são grandes e metacêntricos, o cromossomo 4 é um cromossomo acrocêntrico muito pequeno. Nas glândulas salivares, os cromossomos são muito grandes (ver *Capítulo 11*), o que torna a *Drosophila* um excelente material para estudos genéticos. Em 2000, foi sequenciado o genoma completo da *D. melanogaster*, seguido pelo sequenciamento do genoma da *D. pseudoobscura* em 2005 e de mais 10 genomas de *Drosophila* em 2007. *Drosophila* ainda é um dos organismos-modelo da genética mais versáteis e úteis de todos. ■



Bactéria *Escherichia coli*

O organismo procariótico mais estudado e uma das mais bem caracterizadas geneticamente de todas as espécies é a bactéria *Escherichia coli*. Embora algumas cepas de *E. coli* sejam tóxicas e provoquem doenças, a maioria é benigna e reside naturalmente no sistema digestório dos humanos e outros animais homeotérmicos. *E. coli* foi descrita pela primeira vez por Theodore Escherich em 1885, mas, por muitos anos, acreditava-se que todas as bactérias se reproduzissem por reprodução assexuada e que os cruzamentos genéticos fossem impossíveis. Em 1946, Joshua Lederberg e Edward Tatum demonstraram que *E. coli* sofre um tipo de reprodução sexuada, e esse achado iniciou o uso desta bactéria como organismo-modelo da genética. Um ano depois, Lederberg publicou o primeiro mapa genético da *E. coli* com base nas frequências de recombinação e em 1952, William Hays mostrou que o cruzamento entre as bactérias é assimétrico, com uma bactéria servindo como doadora genética e a outra como receptora genética.

Vantagens de *E. coli* como organismo-modelo da genética. *Escherichia coli* é uma das verdadeiras molas da genética; sua vantagem dupla é a rápida reprodução e o tamanho pequeno. Sob condições ideais, este organismo pode se reproduzir a cada 20 minutos e em apenas 7 horas, uma única célula bacteriana pode dar origem a mais de 2 milhões de descendentes. Uma das importâncias da reprodução rápida é o imenso número de células que podem crescer rapidamente e então até mutações muito raras aparecerão em um curto período de tempo. Consequentemente, numerosas mutações em *E. coli*, que afetam tudo, desde a aparência da colônia à resistência a fármacos, foram isoladas e caracterizadas.

Escherichia coli é fácil de cultivar no laboratório em meio líquido (ver Figura 9.1 A) ou em meio sólido em placas de Petri (ver Figura 9.1 B). Na cultura líquida, *E. coli* cresce até uma concentração de um bilhão de células por mililitro, e trilhões de células bacterianas podem crescer facilmente em um único tubo de teste. Quando as células de *E. coli* são diluídas e espalhadas no meio sólido de uma placa de Petri, as bactérias individuais se reproduzem assexuadamente, dando origem a uma massa concentrada de 10 milhões a 100 milhões de células geneticamente idênticas, chamada de colônia. Essa formação de colônia facilita o isolamento genético das cepas puras da bactéria.

Genoma de *E. coli*. O genoma de *E. coli* é um único cromossomo e, comparado com os cromossomos dos humanos, camundongos, plantas e outros organismos multicelulares, é relativamente pequeno, com 4.638.858 pares de bases. Se fosse esticada, a molécula de DNA no único cromossomo de *E. coli* teria 1,6 mm de comprimento, quase mil vezes o comprimento da célula da *E. coli* dentro da qual ela reside (ver Figura 11.1). Para acomodar esta imensa quantidade de DNA dentro dos limites de uma única célula, o cromossomo de *E. coli* está altamente enroscado e condensado. A informação no cromossomo da bactéria também está compacta, tendo pouco DNA não codificador entre e dentro dos genes e com poucas sequências para as quais existe mais de uma cópia. O genoma de *E. coli* tem em torno de 4.300 genes, muitos dos quais não têm função conhecida. Estes “genes órfãos” podem ser importantes na adaptação a ambientes incomuns, coordenar as vias metabólicas, organizar o cromossomo ou se comunicar com outras células bacterianas. O genoma haploide de *E. coli* facilita isolar as mutações porque não existem genes dominantes no mesmo *loci* para suprimir e mascarar as mutações recessivas.

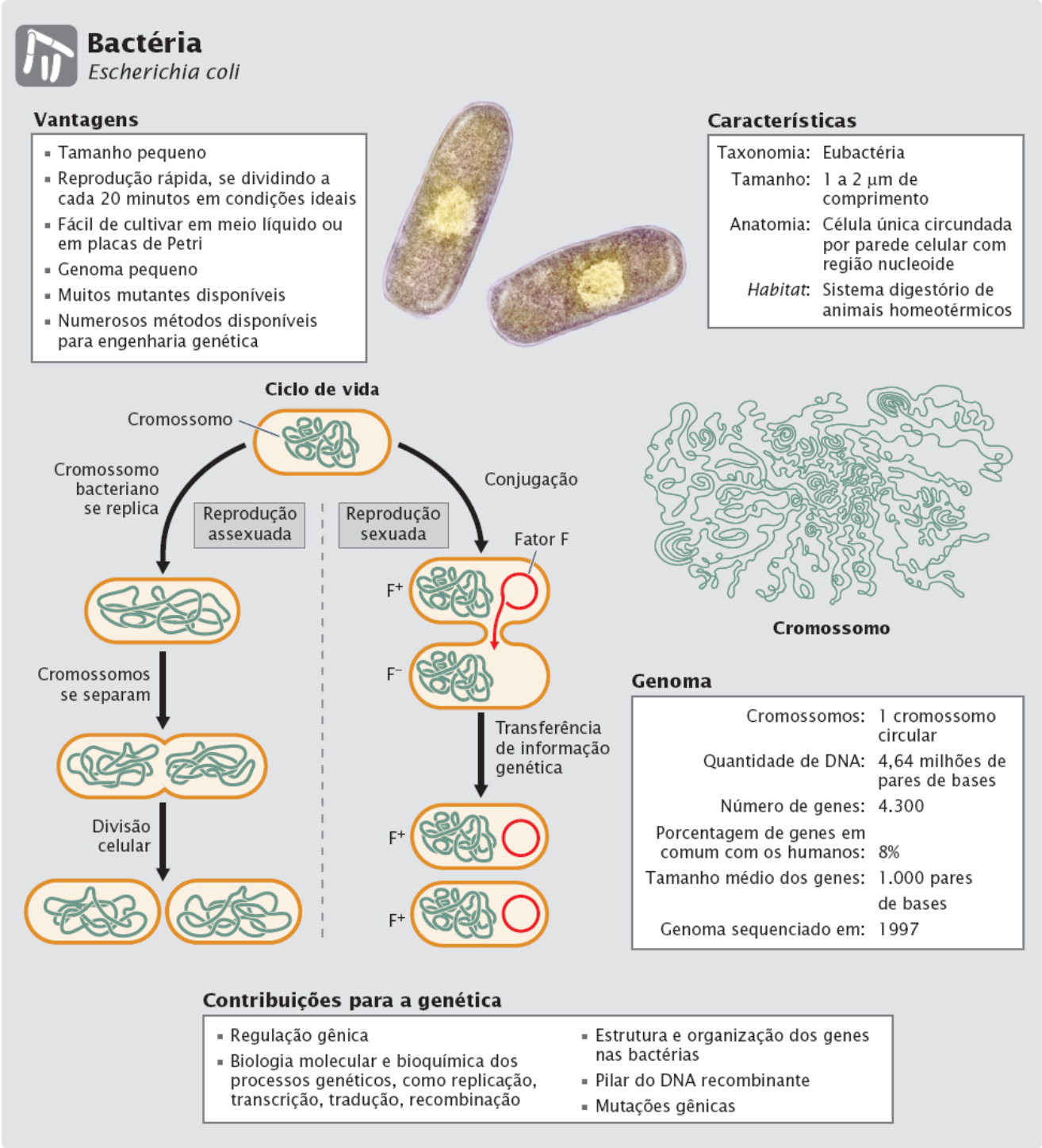
Ciclo de vida de *E. coli*. *E. coli* do tipo selvagem é prototrófica e pode crescer em meio mínimo com apenas glicose e alguns sais inorgânicos. Sob a maioria das condições, esta bactéria se divide em cerca de uma hora, contudo, em meios enriquecidos com açúcares e aminoácidos, ela se dividirá a cada 20 minutos. Normalmente, ela se reproduz por fissão binária simples, na qual o único cromossomo de uma bactéria se replica e migra para lados opostos da célula, seguida pela divisão celular, dando origem a duas células-irmãs idênticas (ver Figura 2.4). O cruzamento entre as bactérias, chamado de conjugação, é controlado por genes de fertilidade, normalmente localizados no plasmídio F (ver *Capítulo 9*). Na conjugação, uma bactéria doa material genético para outra, seguido por recombinação genética que integra novos alelos ao cromossomo bacteriano. O material genético também pode ser trocado entre cepas de *E. coli* por meio de transformação e transdução (ver Figura 9.6).

Técnicas genéticas com *E. coli*. *Escherichia coli* é usada em vários sistemas experimentais nos quais os processos genéticos fundamentais são estudados em detalhes. Por exemplo, os sistemas de tradução *in vitro* têm dentro de um tubo de teste todos os componentes necessários para traduzir a informação genética de RNA mensageiro em uma cadeia de polipeptídios. De forma semelhante, os sistemas *in vitro* com componentes das células de *E. coli* possibilitam o estudo e a análise em condições laboratoriais controladas da transcrição, da replicação e da expressão gênica, além de muitas outras funções genéticas importantes.

Escherichia coli também é muito usada na engenharia genética (DNA recombinante, ver *Capítulo 19*). Plasmídios de *E. coli* foram isolados e geneticamente modificados para criar vetores efetivos para transferir genes nas bactérias e células eucarióticas. Novas construções genéticas (sequências de DNA criadas em laboratório) são montadas e clonadas em *E. coli* antes de serem transferidas para outros organismos. Já foram desenvolvidos métodos para introduzir mutações específicas dentro dos genes de

E. coli, e então a análise genética não depende mais do isolamento de mutações aleatórias. Novas sequências de DNA produzidas por DNA recombinante podem ser introduzidas por transformação em cepas especiais de *E. coli* que são particularmente eficientes (competentes) para captar DNA.

Graças a suas incríveis vantagens como um organismo-modelo da genética, *E. coli* tem papel principal em muitas descobertas na genética, incluindo a elucidação do código genético, sondando a natureza da replicação e organizando os mecanismos básicos da regulação gênica. ■



Nematódeo *Caenorhabditis elegans*

Você pode estar se perguntando, o que é um nematódeo e por que ele é um organismo-modelo da genética? Embora seja raro observá-los, os nematódeos estão entre os organismos mais abundantes na Terra, habitando os solos no mundo todo. A maioria vive livremente e não apresenta perigo, mas alguns são parasitas importantes de plantas e animais, incluindo os humanos. Embora o *Caenorhabditis elegans* não tenha importância econômica ou médica, é usado em estudos de genética por causa do seu único plano de corpo, facilidade para cultivar e alta capacidade de reprodução. Introduzida ao estudo da genética

por Sydney Brenner, que formulou planos em 1962 para usar *C. elegans* para a dissecção genética do comportamento, esta espécie fez importantes contribuições para o estudo do desenvolvimento, morte celular, envelhecimento e comportamento.

Vantagens de *C. elegans* como organismo-modelo da genética. Um organismo genético ideal, *C. elegans* é pequeno, fácil de cultivar e produz muitos descendentes. O adulto *C. elegans* tem cerca de 1 mm de comprimento. A maioria dos investigadores cria *C. elegans* em placas de Petri cheias de ágar que são cobertas por tecido de bactérias, que os nematódeos devoram. Milhares de vermes podem ser cultivados com facilidade em um único laboratório. Comparado com a maioria dos animais multicelulares, eles têm um tempo de geração muito curto, cerca de 3 dias à temperatura ambiente. E eles são reprodutores férteis, com uma única fêmea produzindo de 250 a 1.000 ovos fertilizados em 3 a 4 dias.

Outra vantagem do *C. elegans*, em particular para os estudos do desenvolvimento, é que o verme é transparente, permitindo a fácil observação do desenvolvimento interno em todos os estágios. Ele tem uma estrutura corporal única, com um pequeno número constante de células somáticas: 959 células em uma fêmea hermafrodita madura e 1.031 células em um macho maduro.

Ciclo de vida de *C. elegans*. A maioria dos adultos maduros é hermafrodita, com a capacidade de produzir óvulos e espermatozoides e de autofertilização. Alguns poucos são machos, que produzem apenas espermatozoides e cruzam com hermafroditas. As fêmeas têm dois cromossomos X (XX), e os machos têm um único cromossomo X (X0). Desta forma, os hermafroditas que se autofertilizam produzem apenas hermafroditas (com a exceção de alguns machos que surgem a partir da não disjunção dos cromossomos X). Quando os hermafroditas cruzam com os machos, metade dos descendentes é hermafrodita XX e metade é macho X0. Os óvulos são fertilizados internamente, a partir do espermatozoide produzido pelo hermafrodita ou do espermatozoide de um macho.

Os ovos são postos e o desenvolvimento é completado no meio externo. Aproximadamente 14 horas após a fertilização, uma larva sai do ovo e sofre quatro estágios larvais, chamados L1, L2, L3 e L4, que são separados por mudas. A larva L4 sofre uma muda final para produzir um verme adulto. Sob condições laboratoriais normais, os vermes vivem por 2 a 3 semanas.

Genoma de *C. elegans*. Os geneticistas começaram a desenvolver planos em 1989 para sequenciar o genoma de *C. elegans* e a sequência completa do genoma foi obtida em 1998. Comparado com os genomas da maioria dos animais multicelulares, o genoma de *C. elegans*, com 103 milhões de pares de bases de DNA, é pequeno, o que facilita a análise genômica. A disponibilidade da sequência completa do genoma fornece muitas informações sobre a estrutura, função e organização do gene nesta espécie. Por exemplo, o processo de morte celular programada (apoptose, ver *Capítulo 22*), é importante no desenvolvimento e na supressão do câncer. A apoptose em *C. elegans* é muito semelhante à apoptose nos humanos. Com a sequência genômica completa de *C. elegans*, e considerando sua facilidade para manipulação genética, os geneticistas identificaram genes que participam da apoptose, o que aumentou nossa compreensão sobre a apoptose nos humanos e seu papel no câncer.

Técnicas genéticas com *C. elegans*. Os mutágenos químicos são usados rotineiramente para gerar mutações em *C. elegans*, mutações fáceis de ser identificadas e isoladas. A capacidade dos hermafroditas em se autofertilizar significa que podem ser obtidos descendentes homozigotos para mutações recessivas em uma única geração; a existências dos machos significa que podem ser feitos os cruzamentos genéticos.

Os estudos de desenvolvimento são facilitados pelo corpo transparente dos vermes. Como já foi afirmado, *C. elegans* tem um número pequeno e exato de células somáticas. Os pesquisadores que estudam o desenvolvimento de *C. elegans* mapearam meticulosamente a linhagem celular inteira da espécie e então o destino do desenvolvimento de cada célula no corpo adulto pode ser rastreado a partir do óvulo fertilizado da célula única original. Os biólogos do desenvolvimento usam *lasers* para destruir (remover) células específicas em um verme em desenvolvimento e então estudam os efeitos sobre a fisiologia, o desenvolvimento e comportamento.

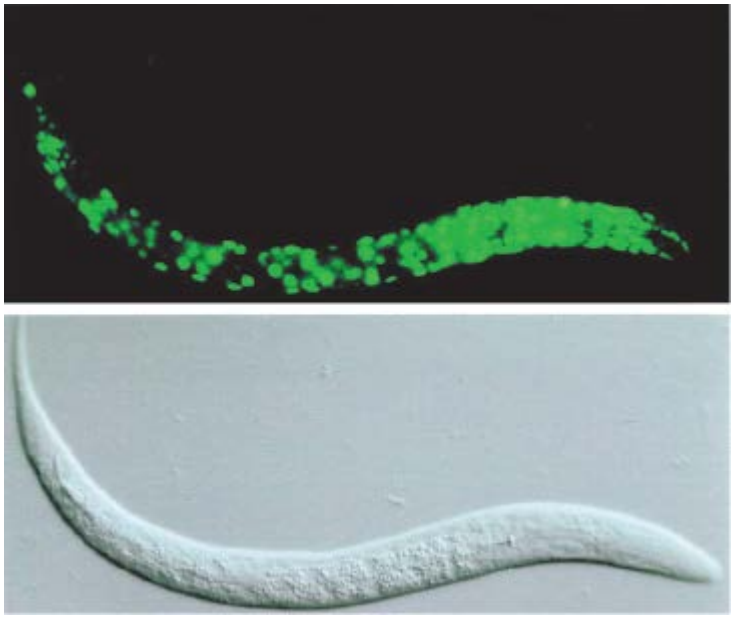
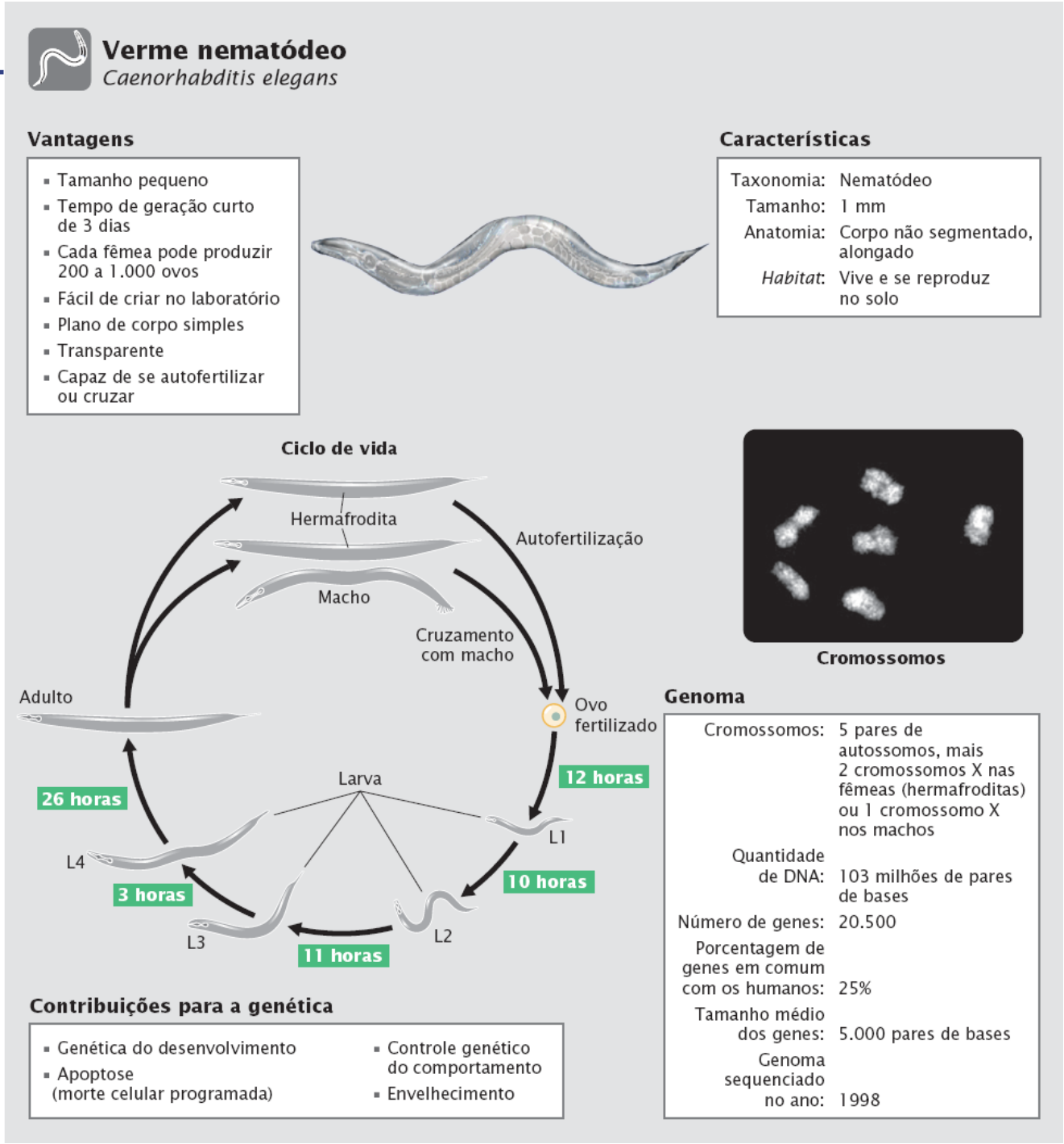


Figura 1 Uma sequência para a proteína fluorescente verde (GFP) é usada para determinar visualmente a expressão dos genes inseridos em *C. elegans* (fotografia inferior). O gene para GFP é injetado no ovário de um verme e é incorporado ao seu genoma. A expressão deste transgene produz GFP, que fluoresce verde (fotografia superior). (Huaqi Jiang, Rong Guo, and Jo Anne Powell-Coffman. The Caenorhabditis elegans hif-1 gene encodes a bHLH-PAS protein that is required for adaptation to hypoxia. PNAS 98: 7916-7921, 2001. ©2001 National Academy of Sciences, U.S.A.)



(Fotografia cortesia de William Goodyer e Monique Zetka.)

A interferência pelo RNA é comprovadamente uma ferramenta efetiva para "desligar" os genes no *C. elegans*. Os geneticistas injetaram cópias de fita dupla de RNA que é complementar a genes específicos; o RNA de fita dupla silencia a expressão destes genes por meio do processo do RNAi. Os vermes podem ser até alimentados por bactérias que foram geneticamente modificadas para expressar o RNA de fita dupla, evitando as dificuldades da microinjeção.

Os vermes transgênicos podem ser produzidos ao injetar DNA no ovário, onde o DNA é incorporado aos ovócitos. Os geneticistas criaram um gene repórter especial que produz a proteína verde fluorescente da medusa (GFP). Quando este gene repórter é injetado no ovário e é inserido no genoma do verme, sua expressão produz GFP, que fluoresce verde, permitindo que a expressão do gene seja facilmente observada (**Figura 1**). ■

Planta *Arabidopsis thaliana*

Parte do trabalho inicial em genética foi feito nas plantas, incluindo as descobertas produtivas de Mendel nas ervilhas assim como aspectos importantes da hereditariedade, mapeamento gênico, genética do cromossomo e herança quantitativa no milho, trigo, feijões e outras plantas. Entretanto, na metade do século 20, muitos geneticistas adotaram bactérias, vírus,

levedura, *Drosophila* e camundongo como modelos genéticos. Como não existia uma boa planta-modelo da genética, as plantas foram negligenciadas, particularmente para o estudo dos processos genéticos moleculares.

Isso mudou na última parte do século 20, com a introdução difundida de um novo organismo-modelo da genética, a planta *Arabidopsis thaliana*. *A. thaliana* foi identificada no século 16, e o primeiro mutante foi reportado em 1873, mas não era comum estudar esta espécie até os primeiros mapas genéticos detalhados surgirem no início dos anos 1980. Atualmente, *Arabidopsis* figura de forma proeminente no estudo da estrutura do genoma, regulação gênica, desenvolvimento e evolução nas plantas e fornece informações básicas importantes sobre a genética das plantas que são aplicadas a espécies economicamente importantes.

Vantagens de *Arabidopsis* como organismo-modelo da genética. A planta *Arabidopsis thaliana* é um membro da família Brassicaceae e cresce como uma erva em muitas partes do mundo. Exceto no seu papel como organismo-modelo da genética, a *Arabidopsis* não tem importância econômica, mas tem várias características que a tornam muito adequada para o estudo da genética. Como um angiosperma, tem as características em comum com outras plantas florescentes, algumas das quais são críticas para o ecossistema ou são fontes importantes de alimento, fibra, material de construção e agentes farmacêuticos. As principais vantagens de *Arabidopsis* são seu tamanho pequeno (altura máxima de 10 a 20 cm), reprodução produtiva e genoma pequeno.

Arabidopsis thaliana termina o desenvolvimento, desde a germinação da semente à produção da semente, em 6 semanas. Seu tamanho pequeno e capacidade de crescer com pouca iluminação a tornam ideal para cultivo no laboratório. Cada planta é capaz de produzir de 10.000 a 40.000 sementes e as sementes têm uma alta taxa de germinação, então podem ser obtidas grandes quantidades de descendentes a partir de cruzamentos genéticos únicos.

Genoma de *Arabidopsis*. Uma vantagem principal para os estudos moleculares é o pequeno genoma de *Arabidopsis*, composto por apenas 125 milhões de pares de bases de DNA em cinco pares de cromossomos, comparado com 2,5 bilhões de pares de bases do DNA no genoma do milho e 16 bilhões de pares de bases no genoma do trigo. O genoma de *A. thaliana* foi completamente sequenciado em 2000, fornecendo informação detalhada sobre a estrutura e organização do gene nesta espécie. Estão disponíveis para estudo muitas variantes de *A. thaliana*, chamada ecotipos, com diferentes formatos, tamanhos e características fisiológicas e sequência de DNA.

Ciclo de vida de *Arabidopsis*. O ciclo de vida de *Arabidopsis* é muito típico da maioria das plantas florescentes. A principal parte vegetativa da planta é diploide, os gametas haploides são produzidos no pólen e ovários. Quando um grão de pólen pousa no estigma de uma flor, um tubo de pólen cresce para o pistilo e ovário. Dois núcleos de espermatozoides haploides contidos em cada pólen viajam pelo tubo do pólen e entram no saco embrionário. Aqui, uma das células do espermatozoide haploide fertiliza o ovo haploide para produzir um zigoto haploide. O outro espermatozoide haploide se funde com dois núcleos haploides para formar o endosperma 3*n*, que fornece tecido que vai nutrir a planta embrionária em crescimento. Os zigotos se desenvolvem dentro das sementes, que são produzidas em uma longa vagem.

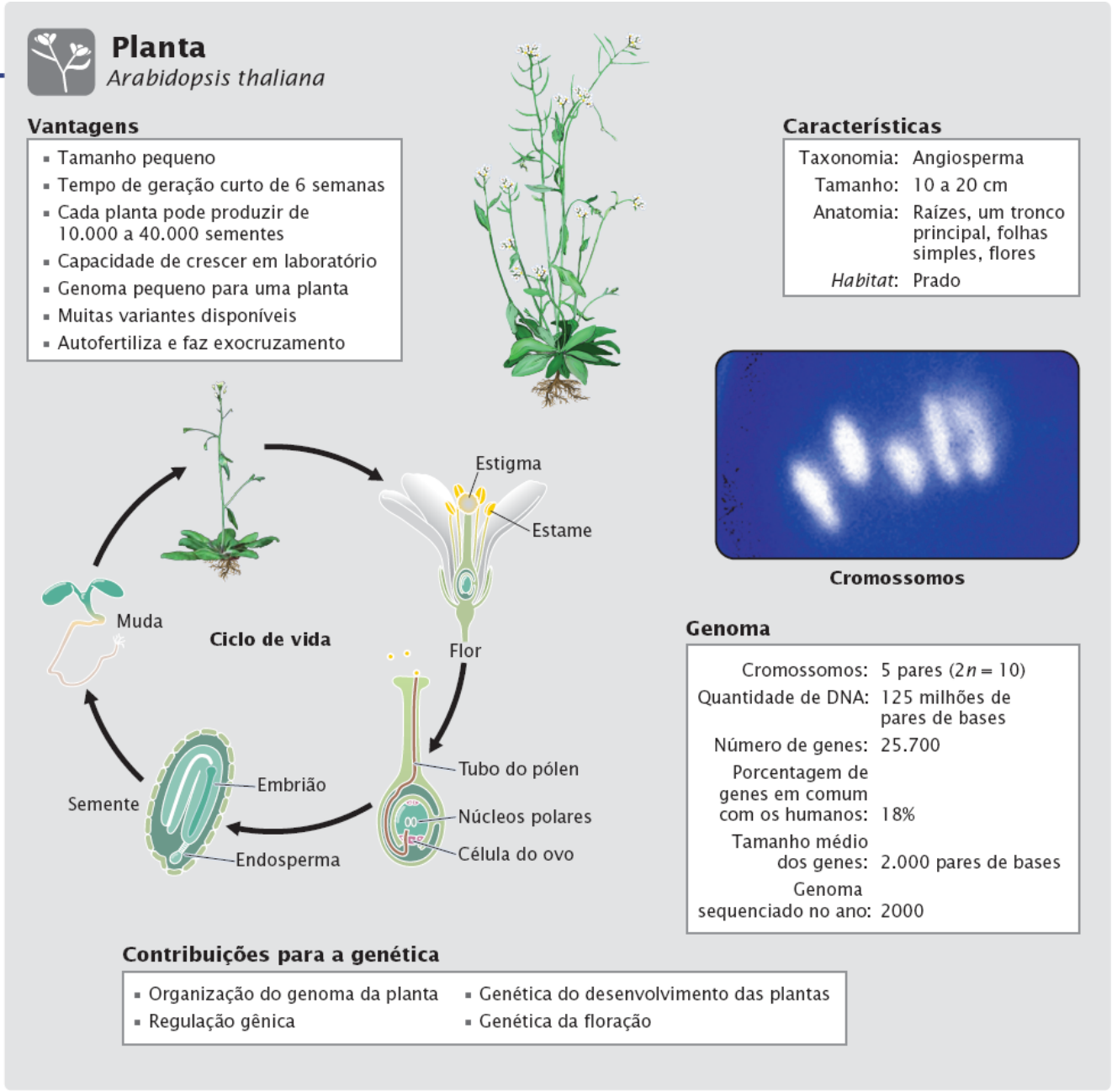
Em condições adequadas, o embrião germina e começa a crescer em uma planta. O tronco cresce para cima e as raízes para baixo, uma roseta compacta de folhas é produzida e sob as condições certas, o tronco aumenta e se diferencia em estruturas florais. Na maturidade, *A. thaliana* é uma planta de crescimento lento com raízes, um tronco principal com galhos que sustentam folhas maduras e pequenas flores brancas nas pontas dos galhos.

Técnicas genéticas com *Arabidopsis*. É comum usar várias técnicas moleculares tradicionais e modernas com *Arabidopsis* e obter vantagens especiais para estudos genéticos. *Arabidopsis* pode se autofertilizar, o que significa que qualquer mutação recessiva que apareça na linhagem germinativa pode ser recuperada nos descendentes imediatos. A fertilização cruzada também é possível ao remover a antera de uma planta e grãos de pólen do estigma de outra planta, essencialmente a mesma técnica usada por Gregor Mendel com as ervilhas (ver Figura 3.4).

Como já foi mencionado, muitas variantes de ocorrência natural de *Arabidopsis* estão disponíveis para estudos e novas mutações podem ser produzidas ao expor suas sementes a mutágenos químicos, radiação ou elementos de transposição que se inserem aleatoriamente nos genes. O grande número de descendentes produzido por *Arabidopsis* facilita o rastreamento de mutações raras.

Os genes de outros organismos podem ser transferidos para *Arabidopsis* pelo plasmídio Ti da bactéria *Agrobacterium tumefaciens*, que naturalmente infecta plantas e transfere o plasmídio Ti para as células da planta (ver *Capítulo 19*). Após a transferência para uma célula da planta, o plasmídio Ti se insere aleatoriamente no DNA da planta que infecta, gerando mutações no DNA da planta em um processo chamado mutagênese insercional. Os geneticistas modificaram o plasmídio Ti para carrear o gene *GUS*, que não tem seu próprio promotor. O gene *GUS* codifica uma enzima que converte um composto incolor (X-Glu) em um corante azul. Como este gene não tem promotor, ele é expresso apenas quando inserido na sequência

codificadora de um gene da planta. Quando isto acontece, a enzima codificada pelo *GUS* é sintetizada e converte X-Glu em um corante azul que cora a célula. Este corante fornece um meio para determinar visualmente o padrão de expressão de um gene que foi interrompido pelo DNA do Ti, produzindo informações sobre a expressão dos genes que sofrem mutação pela mutagênese insercional. ■



(Fotografia cortesia de Anand P Tyagi e Luca Comai.)

Camundongo *Mus musculus*

O camundongo comum doméstico, *Mus musculus*, é um dos mais antigos e mais valiosos materiais para estudo genético. É um excelente organismo genético, pequeno, fértil, fácil de manter, com tempo de geração curto.

Vantagens do camundongo como organismo-modelo da genética. A primeira entre as muitas vantagens que o *Mus musculus* tem como organismo-modelo da genética é sua relação evolutiva próxima aos humanos. Sendo um mamífero, o camundongo é, do ponto de vista genético, comportamental e fisiológico muito mais semelhante aos humanos do que outros organismos usados em estudos genéticos, tornando o camundongo o modelo de escolha para muitos estudos da genética humana e médica. Outras vantagens incluem um curto tempo de geração em comparação com a maioria dos outros mamíferos. *Mus musculus* está bem adaptado à vida no laboratório e pode ser criado e reproduzido em gaiolas que exigem pouco espaço, assim milhares de camundongos podem ser criados dentro dos limites de um pequeno laboratório. Os camundongos têm grandes ninhadas (8 a 10 filhotes) e são dóceis e fáceis de manipular. Finalmente, foram isoladas e estudadas várias mutações em

camundongos criados em cativeiro, fornecendo uma fonte importante de variação para análise genética.

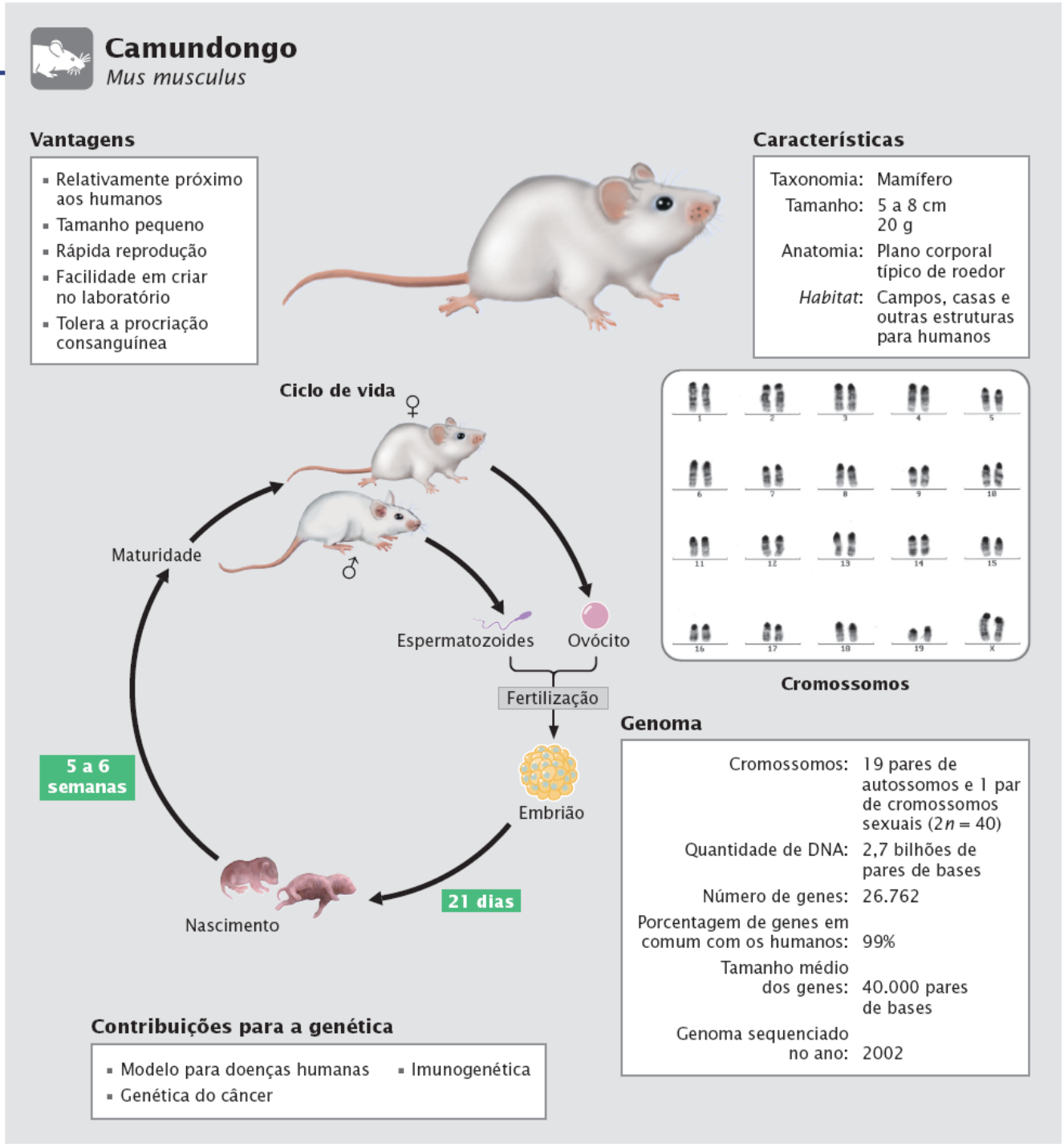
Ciclo de vida do camundongo. A produção de gametas e a reprodução no camundongo são muito semelhantes às dos humanos. As células germinativas diploides nas gônadas sofrem meiose para produzir espermatozoides e ovócitos, como destacado no Capítulo 2. Os camundongos-machos começam a produzir espermatozoides na puberdade e continuam a produzir pelo resto de suas vidas. Começando na puberdade, as fêmeas entram no cio a cada 4 dias. Se ocorrer cruzamento durante o cio, os espermatozoides são depositados na vagina e nadam até o oviduto, onde um espermatozoide penetra na camada externa do ovo e os núcleos do espermatozoide e do óvulo se fundem. Após a fertilização, o embrião diploide se implanta no útero. Em geral a gestação dura 21 dias. Os camundongos alcançam a puberdade em 5 a 6 semanas e vivem por 2 anos. Uma geração completa pode ser obtida em 8 semanas.

Genoma do camundongo. O genoma do camundongo tem cerca de 2,6 bilhões de pares de bases de DNA, que é semelhante ao tamanho do genoma humano. Os camundongos e humanos também têm números de genes semelhantes. Para a maioria dos genes humanos, existem genes homólogos no camundongo. Uma ferramenta importante para determinar a função de um gene desconhecido nos humanos é pesquisar um gene homólogo cuja função já tenha sido determinada no camundongo. Além disso, as relações de ligação de muitos genes de camundongos são semelhantes às dos humanos, e as relações de ligação dos genes nos camundongos fornecem dicas importantes sobre essas relações entre os genes nos humanos. O genoma do camundongo é distribuído em 19 pares de autossomos e um par de cromossomos sexuais.

Técnicas genéticas com o camundongo. Foram desenvolvidas várias técnicas versáteis para usar no camundongo. Elas incluem a criação de camundongos transgênicos com a injeção de DNA em um embrião de camundongo, a capacidade de romper genes específicos por meio da criação de camundongos *knockout* e a capacidade de inserir sequências específicas em *loci* específicos (ver *Capítulo 19*). Estas técnicas possibilitaram manipular o ciclo reprodutivo do camundongo, incluindo a capacidade de induzir a ovulação com hormônios, isolar ovócitos não fertilizados do ovário e implantar embriões fertilizados de volta no útero de uma mãe substituta. A capacidade de criar camundongos transgênicos, *knockout* e *knock-in* facilitou muito o estudo da genética humana, e estas técnicas ilustram a força do camundongo como um organismo-modelo da genética.

As células de camundongo e de seres humanos podem ser fundidas, permitindo as técnicas de hibridização de célula somática (ver *Capítulo 7*) que são muito usadas para designar os genes humanos a cromossomos específicos. Os camundongos também toleram bem a endogamia, e é fácil criar cepas consanguíneas com o cruzamento de irmão-irmã. Membros de uma cepa consanguínea são geneticamente muito semelhantes ou idênticos, permitindo que os pesquisadores examinem os efeitos dos fatores ambientais sobre um traço.

O uso dos camundongos como um organismo-modelo da genética levou a importantes descobertas genéticas. No início do século 20, os camundongos eram usados para estudar a base genética da variação da cor do pelo nos mamíferos. Nos últimos anos, eles se tornaram uma parte importante na pesquisa sobre a base genética do câncer e os potenciais carcinógenos são testados em camundongos. Os camundongos são usados para estudar genes que influenciam o desenvolvimento dos mamíferos, incluindo mutações que produzem defeitos de nascimento nos humanos. Foram criados vários modelos de camundongos de doenças humanas específicas – em alguns casos, ao isolar e cruzar camundongos com mutações naturais e, em outros casos, ao usar técnicas de *knockout* e *knock-in* e modificar genes específicos. ■



(Fotografia cortesia de Ellen C. Akeson e Muriel T. Davisson, The Jackson Laboratory, Bar Harton, Maine.)

Levedura *Saccharomyces cerevisiae*

A levedura comum de pão (*Saccharomyces cerevisiae*) é amplamente adotada como um sistema-modelo único para o estudo da genética dos eucariotos. Usada há muito tempo para preparar pão e cerveja, a levedura vem sendo usada recentemente para a produção de biocombustíveis. Louis Pasteur identificou a *S. cerevisiae* como o microrganismo responsável pela fermentação em 1857, e seu uso na análise genética começou em 1935. Como material de extensivos estudos em genética clássica por muitos anos, os genes da levedura são bem conhecidos e caracterizados. Ao mesmo tempo, a natureza unicelular da levedura a torna adequada para as técnicas moleculares desenvolvidas para bactérias. Assim, a levedura combina a genética clássica e a biologia molecular para fornecer um modelo versátil para o estudo de sistemas genéticos em eucariotos.

Vantagens da levedura como organismo-modelo da genética. A grande vantagem da levedura é que ela não é apenas um organismo eucariótico, com sistemas genético e celular semelhantes aos de outros eucariotos mais complexos, como os humanos, mas também é unicelular, com muitas das vantagens de manipulação dos sistemas bacterianos. Como as bactérias, as

células de levedura precisam de pouco espaço e numerosas células podem ser cultivadas com facilidade e sem custo no laboratório.

A levedura existe nas formas diploide e haploide. Quando haploide, as células têm apenas um único alelo em cada *locus*, o que significa que o alelo será expresso no fenótipo; diferente do diploide, não existe dominância no qual alguns alelos mascaram a expressão de outros. Portanto, os alelos recessivos podem ser facilmente identificados nas células haploides e então as interações dos alelos podem ser examinadas nas células diploides.

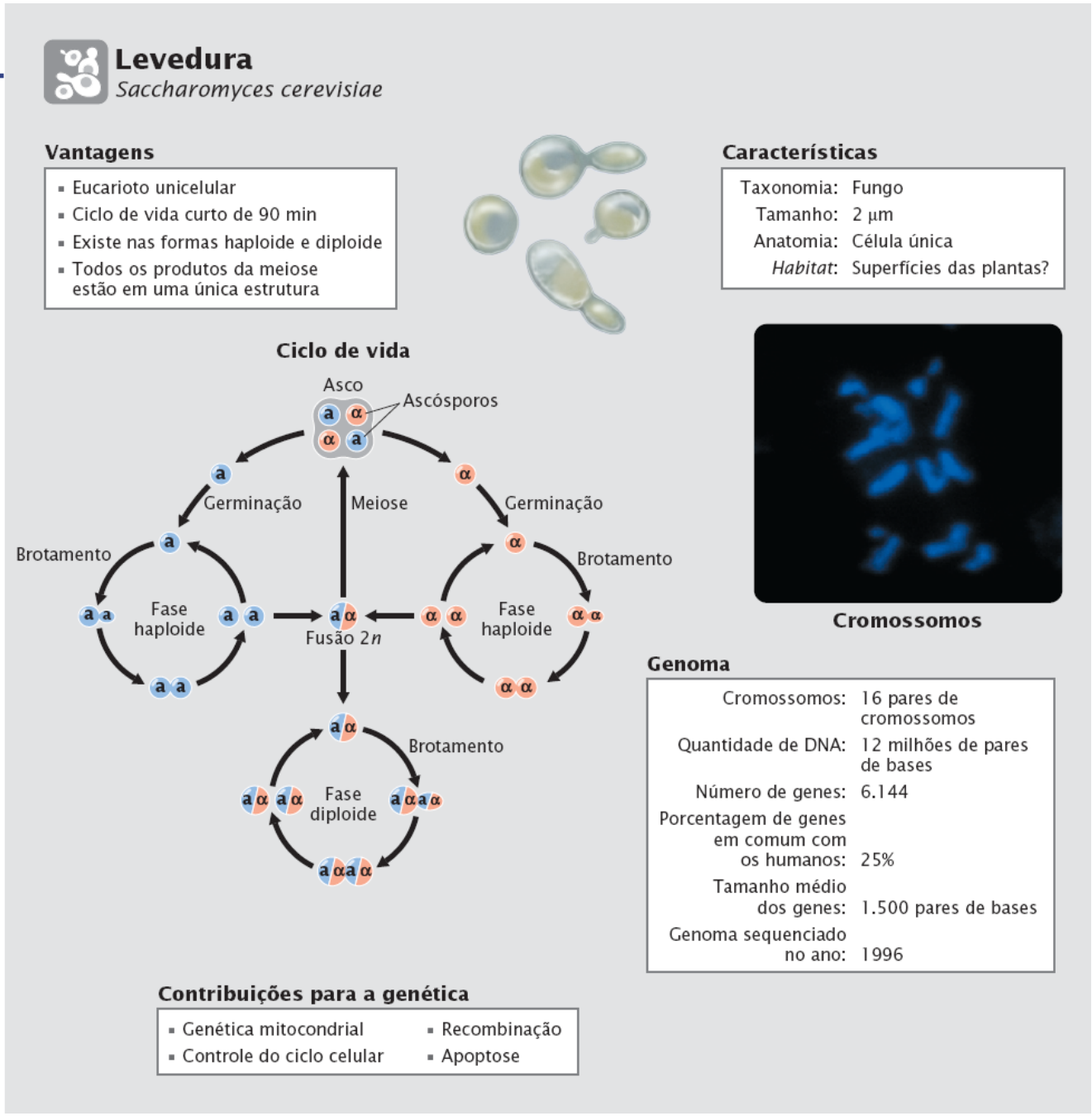
Outra característica que torna a levedura um sistema-modelo genético versátil é que, subsequentemente à meiose, todos os produtos de uma divisão meiótica estão presentes em uma única estrutura chamada asco (ver a próxima subseção) e permanecem separados dos produtos de outras divisões meióticas. As quatro células produzidas por uma única divisão meiótica são chamadas tétrades. Na maioria dos organismos, os produtos de diferentes divisões meióticas se misturam e é impossível identificar os resultados de uma única divisão meiótica. Por exemplo, se formos isolar quatro espermatozoides do testículo de um camundongo, é muito improvável que todas as quatro células tenham sido produzidas a partir da mesma divisão meiótica. As tétrades separadas na levedura possibilitam que observemos diretamente os efeitos das divisões meióticas individuais sobre os tipos de gametas produzidos e mais facilmente identificamos os eventos de *crossing over*. A análise genética de uma tétrade é chamada de análise de tétrade.

A levedura tem sido submetida à extensiva análise genética e foram identificados milhares de mutantes. Além disso, muitas técnicas moleculares versáteis desenvolvidas para manipular sequências genéticas nas bactérias foram adaptadas para uso na levedura. Ainda, apesar de uma estrutura unicelular e da facilidade de manipulação, as células da levedura têm muitos dos mesmos genes encontrados nos humanos e outros eucariotos multicelulares complexos e muitos destes genes têm funções idênticas ou semelhantes nestes eucariotos. Assim, o estudo da genética das células da levedura contribui para nossa compreensão de outros organismos eucarióticos mais complexos, incluindo os humanos.

Ciclo de vida da levedura. Como foi afirmado anteriormente, *Saccharomyces cerevisiae* pode existir como células haploides ou diploides. As células haploides existem quando a levedura está carente de nutrientes e se divide por mitose, produzindo células-irmãs haploides idênticas, por germinação. As células da levedura também podem apresentar reprodução sexuada. Existem dois tipos de acasalamentos, a e α; as células haploides de diferentes tipos de acasalamento se fundem e então sofrem fusão nuclear para criar uma célula diploide. A célula diploide é capaz de germinar por mitose para produzir células diploides geneticamente idênticas. A falta de nutrientes induz as células diploides a sofrerem meiose, resultando em quatro núcleos haploides, que se separam em diferentes células, produzindo esporos haploides. Como os quatro produtos da meiose (uma tétrade) estão presos em uma estrutura comum, o asco, todos os produtos de uma única meiose podem ser isolados (análise de tétrade).

Genoma da levedura. *Saccharomyces cerevisiae* tem 16 pares de cromossomos eucarióticos típicos. A taxa de recombinação é alta, dando à levedura um mapa genético relativamente longo se comparado com o de outros organismos. O genoma de *S. cerevisiae* tem 12 milhões de pares de bases, mais os 2 a 3 milhões de pares de bases dos genes rRNA. Em 1996, *S. cerevisiae* foi o primeiro organismo eucariótico cujo genoma foi completamente sequenciado.

Técnicas genéticas com a levedura. Uma vantagem da levedura para os pesquisadores é o uso de plasmídios para transferir genes ou sequências de DNA de interesse para as células. As células da levedura naturalmente têm um plasmídio circular, chamado 2μ, com 6.300 pb de comprimento e é transmitido para as células-filhas na mitose e meiose. Este plasmídio tem uma origem de replicação reconhecida pelo sistema de replicação da levedura, então ele se replica de forma autônoma na célula. O plasmídio 2μ foi submetido à manipulação para fornecer um vetor eficiente para transferir genes para a levedura. Em outros casos, os plasmídios bacterianos são adaptados para serem usados na levedura. Alguns deles sofrem recombinação homóloga com o cromossomo da levedura, transferindo suas sequências para o cromossomo da levedura. Os vetores de transferência (*shuttle vectors*), que podem ser propagados nas bactérias e leveduras, são particularmente eficientes. Tais vetores possibilitam a criação e a manipulação de sequências de genes nas bactérias, nos locais que dispõem de técnicas frequentemente mais efetivas para manipulação e seleção genética, com posterior transferência das sequências de genes para as células da levedura, onde sua função pode ser testada.



(Fotografia cortesia de Mara Stewart e Dean Dawson, Department of Microbiology and Molecular Biology, Sackler School of Biomedical Sciences, Tufts University.)

Os plasmídios são limitados no tamanho dos fragmentos de DNA que eles conseguem carrear (ver *Capítulo 19*). Os cromossomos artificiais de levedura (YACs), fragmentos de DNA modificados que têm sequências centroméricas e teloméricas e se separam como cromossomos na meiose e na mitose, superam esta limitação; os YACs conseguem lidar com fragmentos de DNA até centenas de milhares de pares de bases. ■

Glossário

Acasalamento preferencial negativo. Acasalamento entre indivíduos diferentes que ocorre mais frequentemente do que esperado com base no acaso.

Acasalamento preferencial positivo. Acasalamento entre indivíduos semelhantes que ocorre mais frequentemente do que esperado com base no acaso.

Acentuador. Do inglês, *enhancer*. Sequência que estimula a transcrição máxima de genes distantes; afeta apenas os genes na mesma molécula de DNA (ação cis), tem pequenas sequências consenso, não está fixa na relação ao sítio de início da transcrição, consegue estimular quase qualquer promotor na sua vizinhança e pode ser *upstream* ou *downstream* do gene. A função de um acentuador é independente da orientação da sequência.

Aconselhamento genético. Processo de orientação que tenta auxiliar os pacientes e familiares a lidar com todos os aspectos de uma condição genética.

Adenina (A). Base purina no DNA e RNA.

A-DNA. Estrutura helicoidal dextrógira (gira para a direita) do DNA que existe quando há pouca água no ambiente.

Agente intercalante. Substância química que tem aproximadamente o mesmo tamanho de um nucleotídio e se localiza entre bases adjacentes no DNA, distorcendo a estrutura tridimensional da hélice e levando a inserções e deleções de nucleotídio único na replicação.

Alça D. Região do DNA mitocondrial que tem uma origem de replicação e promotores; é deslocada durante o início da replicação, daí o nome alça de deslocamento ou alça D.

Alelo. Uma ou mais formas alternativas de um gene.

Alelo letal. Leva à morte de um organismo, muitas vezes no início do desenvolvimento, e então o organismo não aparece entre os descendentes de um cruzamento genético. Os alelos letais recessivos matam organismos homozigotos para o alelo; os alelos letais dominantes matam tanto heterozigotos como homozigotos.

Alelo sensível à temperatura. Expresso apenas em determinadas temperaturas.

Alelos múltiplos. Em um grupo, existência de indivíduos de mais de dois alelos em um *locus*. Entretanto, cada membro do grupo tem apenas dois dos possíveis alelos.

Alopoliploidia. Condição na qual os conjuntos de cromossomos de um indivíduo poliploide com mais de dois conjuntos haploides são derivados de duas ou mais espécies.

Alterações epigenéticas. Modificações estáveis da estrutura da cromatina que podem ser transmitidas para outras células ou para um organismo individual. (Ver também *Epigenética*.)

Alternância de gerações. Ciclo de vida complexo em plantas que alterna entre o estágio esporófito diploide e o estágio gametófito haploide.

Aminoácido. Unidades repetidas de proteínas; composto por um grupo amino, um grupo carboxila, um átomo de hidrogênio e um grupo R variável.

Aminoacil-tRNA sintetase. Enzima que liga aminoácido a um tRNA. Cada aminoacil-tRNA sintetase é específica para um determinado aminoácido.

Amniocentese. Procedimento pré-natal para obter uma amostra do líquido amniótico de uma gestante. Uma longa agulha estéril é inserida através da parede abdominal no saco amniótico para obter o líquido; depois é feita a análise genética.

Amostra. Subconjunto usado para descrever uma população.

Amostragem das vilosidades coriônicas (CVS). Procedimento pré-natal no qual um pequeno pedaço do cório (a camada externa da placenta) é removido da gestante. Um cateter é inserido através da vagina e do colo uterino e alcança o útero. É aplicada sucção para retirar a amostra.

Anáfase. Estágio da mitose no qual as cromátides se separam e se movem para os polos do fuso.

Anáfase I. Um dos estágios da meiose I. Na anáfase I, os cromossomos homólogos se separam e se movem para os polos do fuso.

Anáfase II. Um dos estágios da meiose II. Na anáfase II, as cromátides se separam e se movem para os polos do fuso.

Anagênese. Mudança evolutiva em uma única linhagem.

Análise de ligação. Mapeamento gênico que se baseia na detecção da ligação física entre os genes, como medida pela taxa de recombinação nos descendentes de um cruzamento.

Análise de tétrade. Análise genética de uma tétrade, os produtos de uma única meiose.

Análogo da base. Substância química com estrutura semelhante à de uma das quatro bases padrões do DNA e que pode ser incorporada a moléculas de DNA recém-sintetizadas na replicação.

Aneuploidia. Mudança no número de cromossomos a partir do tipo selvagem; mais frequentemente, aumento ou redução de um ou dois cromossomos.

Anfidiploidia. Tipo de alopoliploidia em que dois genomas diploides diferentes são combinados de modo que cada cromossomo tem um, e apenas um, parceiro homólogo, e o genoma é funcionalmente diploide.

Antecipação. Gravidade crescente ou idade precoce do início de um traço genético nas gerações seguintes. Por exemplo, as manifestações de uma doença genética podem se tornar mais graves à medida que o traço é transmitido de geração para geração.

Anticódon. Sequência de três nucleotídios no RNA transportador que pareia com o códon correspondente no RNA mensageiro na tradução.

Anticorpo. Produzido por linfócitos B, uma proteína que circula no sangue e em outros líquidos corporais. Um anticorpo se liga a um antígeno específico e marca o antígeno para ser destruído por uma célula fagocitária.

Antígeno. Substância reconhecida pelo sistema imunológico e que provoca uma resposta imunológica.

Antígeno do complexo principal de histocompatibilidade (MHC). Do inglês, *major histocompatibility complex*. Pertence a um grupo grande e diverso de antígenos encontrados na superfície das células que marcam essas células como próprias; codificado por um grande grupo de genes conhecidos como complexo principal de histocompatibilidade. Os linfócitos T se ligam simultaneamente a antígenos estranhos e do MHC.

Antiparalela. Refere-se à característica da dupla-hélice do DNA na qual as duas cadeias de polinucleotídios correm em sentidos opostos.

Antiterminadora. Proteína ou sequência de DNA que inibe o término da transcrição.

Aparato basal de transcrição. Complexo de fatores de transcrição, RNA polimerase e outras proteínas que se unem ao promotor e são capazes de iniciar níveis mínimos de transcrição.

Apoptose. Morte celular programada, na qual uma célula degrada seu próprio DNA, o núcleo e o citoplasma encolhem e a célula sofre fagocitose por outras células sem extravazamento do seu conteúdo.

Aptidão. Sucesso reprodutivo de um genótipo comparado com o de outros genótipos em uma população.

Archaea. Uma das três divisões principais da vida. Archea inclui organismos unicelulares com células procarióticas.

Árvore filogenética com raiz. Árvore filogenética na qual um nó interno representa o ancestral comum de todos os outros organismos (nós) na árvore. Em uma árvore com raiz, todos os organismos descritos têm um ancestral em comum.

Árvore filogenética. Representação gráfica das conexões evolutivas entre organismos ou genes.

Árvore gênica. Árvore filogenética que representa correlações evolutivas em um conjunto de genes.

Atenuação. Tipo de regulação de gene em alguns óperons bacterianos, nos quais a transcrição é iniciada mas termina de forma prematura, antes da transcrição dos genes estruturais.

Atenuador. Estrutura secundária que se forma na região 5′ não traduzida de alguns óperons e leva ao término prematuro da transcrição.

Ativador. Ver *Proteína ativadora de transcrição*.

Ato da não discriminação na informação genética (GINA). Do inglês, Genetic Information Nondiscrimination Act. Lei dos EUA que proíbe o uso de informações genéticas pelos profissionais de saúde para tomada de decisões sobre a cobertura de assistência médica e valores; evita que os empregadores usem informações genéticas nas decisões sobre emprego; também evita que os profissionais de saúde e empregadores solicitem ou exijam que uma pessoa faça um exame genético.

Autopoliploidia. Condição na qual todos os conjuntos de cromossomos de um indivíduo poliploide com mais de dois conjuntos haploides são derivados de uma única espécie.

Autossomo. Cromossomo que é o mesmo em machos e fêmeas; cromossomo não sexual.

Bactéria auxotrófica. Uma bactéria ou fungo com mutação nutricional compromete sua capacidade de sintetizar uma molécula

biológica essencial; não cresce no meio de cultura mínimo, mas consegue crescer no meio de cultura mínimo ao qual foi adicionada a molécula biológica que ela não consegue sintetizar.

Bactéria prototrófica. Um tipo selvagem de bactéria que consegue usar uma fonte de carbono, elementos essenciais como nitrogênio e fósforo, algumas vitaminas e outros íons e nutrientes necessários para sintetizar todos os compostos de que ela precisa para o crescimento e a reprodução.

Bacteriófago. Vírus que infecta células bacterianas.

Base. Ver *Base nitrogenada.*

Base modificada. Base rara encontrada em algumas moléculas de RNA. Tais bases são formas modificadas das bases padrão (adenina, guanina, citosina e uracila).

Base nitrogenada. Base contendo nitrogênio que é uma das três partes de um nucleotídio.

B-DNA. Estrutura helicoidal do DNA girando para a direita que existe quando há água em abundância; a estrutura secundária descrita por Watson e Crick e provavelmente a estrutura de DNA mais comum nas células.

Biblioteca de cDNA (DNA complementar). Coleção de colônias de bactérias ou colônias de fagos com fragmentos de DNA que foram produzidos por transcrição reversa do mRNA celular.

Biblioteca de DNA. Coleção de colônias bacterianas com todos os fragmentos de DNA de uma fonte.

Biblioteca genômica. Coleção de colônias de bactérias ou fagos com fragmentos de DNA que consistem no genoma inteiro de um organismo.

Bioinformática. Síntese de biologia molecular e ciência computacional que desenvolve bancos de dados e ferramentas computacionais para armazenar, restaurar e analisar dados de sequências de ácidos nucleicos e proteínas.

Biotecnologia. Uso de processos biológicos, em particular a genética molecular e a tecnologia do DNA recombinante, para produzir produtos de interesse comercial.

Bivalente. Refere-se a um par emparelhado de cromossomos homólogos.

Bolha de replicação. Segmento de uma molécula de DNA que é desenrolada e sofre replicação.

Bolha de transcrição. Região de uma molécula de DNA que é desenrolada para expor um molde de fita única, que está sendo transcrito em RNA.

Braço aceptor. O braço no tRNA em que um aminoácido se prende.

Cadeia de polinucleotídios. Série de nucleotídios unidos por ligações fosfodiéster.

Caminhada cromossômica. Método empregado para localizar um gene ao usar clones genômicos parcialmente sobrepostos para mover em etapas a partir do gene de interesse ligado, previamente clonado.

Camundongo knock-in. Camundongo que carreia uma sequência estranha inserida em uma localização específica dos cromossomos.

Camundongo knockout. Camundongo no qual um gene normal foi desativado (“nocauteado”).

Camundongo transgênico. Camundongo cujo genoma contém um gene estranho ou genes adicionados ao empregar os métodos de DNA recombinante.

Cap 5′. Extremidade 5′ modificada do mRNA eucariótico, consistindo de um nucleotídio extra (metilado) e metilação da posição 2′ do açúcar ribose em um ou mais nucleotídios subsequentes; tem papel na ligação do ribossomo ao mRNA e afeta a estabilidade do mRNA e a remoção dos íntrons.

Captura por afinidade. Uso de um anticorpo para capturar uma proteína a partir de uma mistura complexa de proteínas. A proteína capturada estará ligada a qualquer outra proteína com a qual interage, e pode então ser analisada por meio de espectrometria de massa para identificar essas proteínas.

Característica com limiar. Característica descontínua cuja expressão depende da suscetibilidade subjacente que varia de forma contínua.

Característica contínua. Apresenta um grande número de possíveis fenótipos que não são facilmente diferenciados, como o peso humano.

Característica descontínua. Exibe apenas alguns fenótipos de fácil distinção. Um exemplo é o formato da semente: liso ou rugoso.

Característica influenciada pelo sexo. Característica codificada por genes autossômicos que são expressos mais facilmente em

um sexo. Por exemplo, um gene autossômico dominante pode ter maior penetrância em machos do que em fêmeas ou um gene autossômico pode ser dominante nos machos e recessivo nas fêmeas.

Característica ligada ao sexo. Característica determinada por um gene ou genes nos cromossomos sexuais.

Característica ligada ao X. Característica determinada por um gene ou genes no cromossomo X.

Característica limitada pelo sexo. Característica codificada por genes autossômicos que são expressos em apenas um sexo. Machos e fêmeas carreiam genes para as características limitadas pelo sexo, mas estas aparecem em apenas um dos sexos.

Característica merística. Característica cujo fenótipo varia em números inteiros, como o número de vértebras.

Característica multifatorial. Determinada por múltiplos genes e fatores ambientais.

Característica poligênica. Codificada por genes em muitos *loci*.

Característica quantitativa. Característica contínua; apresenta um vasto número de possíveis fenótipos, que devem ser descritos por uma medida quantitativa.

Características ligadas ao Y. Característica determinada por um gene ou genes no cromossomo Y.

Cariótipo. Imagem de um conjunto completo dos cromossomos na metáfase de um organismo.

Caspase. Enzima que quebra outras proteínas e regula a apoptose. Cada caspase é sintetizada como um grande precursor inativo (uma pró-caspase) que é ativado por clivagem, frequentemente por outra caspase.

Cauda poli(A). Sequência de nucleotídios de adenina adicionada à extremidade 3′ de um mRNA eucariótico após a transcrição.

Célula competente. Célula capaz de captar o DNA a partir do seu ambiente (capaz de ser transformada).

Célula de memória. Linfócito de vida longa entre o clone de células geradas quando um antígeno estranho é encontrado. Se o mesmo antígeno é encontrado mais uma vez, as células de memória se dividem rapidamente e dão origem a outro clone de células específico para aquele antígeno em questão.

Células-tronco. Células indiferenciadas capazes de formar todo tipo de célula em um organismo.

Células-tronco pluripotentes induzidas (iPSCs). Células adultas artificialmente induzidas a se desdiferenciar e reverter para células-tronco pluripotentes capazes de se tornar muitos tipos de células.

CentiMorgan. Outro nome para unidade de mapa.

Centrifugação por gradiente de densidade de equilíbrio. Método usado para separar moléculas ou organelas de diferente densidade por centrifugação.

Centríolo. Organela citoplasmática composta por microtúbulos; presente em cada polo do fuso nas células animais.

Centrômero. Região contraída de um cromossomo que cora menos fortemente que o resto do cromossomo; região onde os microtúbulos do fuso se prendem ao cromossomo.

Centrossomo. Estrutura a partir da qual o aparato do fuso se desenvolve: contém o centríolo.

Cerne da enzima. Parte da RNA polimerase bacteriana que, durante a transcrição, catalisa o alongamento da molécula de RNA com a adição de nucleotídios RNA; tem quatro subunidades: duas cópias de alfa (α), uma única cópia do beta (β) e uma única cópia de β′ ($\alpha_2\beta\beta'$).

Cerne do promotor. Localizado imediatamente *upstream* do promotor eucariótico, consiste em sequências de DNA às quais o aparelho basal de transcrição se liga.

Chaperona molecular. Molécula que auxilia na dobra adequada de outra molécula.

Ciclina. Uma proteína-chave no controle do ciclo celular; combina-se com uma quinase dependente de ciclina (CDK). Os níveis de ciclina sobem e descem no curso do ciclo celular.

Ciclo celular. Estágios pelos quais uma célula passa de uma divisão celular para a próxima.

Ciclo lisogênico. Ciclo de vida de um bacteriófago no qual os genes do fago primeiro se integram ao cromossomo bacteriano e não são imediatamente transcritos e traduzidos.

Ciclo lítico. Ciclo de vida de um bacteriófago no qual os genes do fago são transcritos e traduzidos, novas partículas de fago são produzidas e a célula hospedeira é rompida.

Cinetócoro. Conjunto de proteínas que se montam no centrômero, fornecendo o ponto de ligação para os microtúbulos do fuso.

Citocinese. Processo pelo qual o citoplasma de uma célula se divide.

Citosina (C). Base pirimidina no DNA e RNA.

Cladogênese. Evolução na qual uma linhagem é dividida em duas.

Clonagem de genes. Inserção de fragmentos de DNA em bactérias de modo que os fragmentos serão estáveis e copiados pelas bactérias.

Clonagem posicional. Método que permite o isolamento e a identificação de um gene ao examinar a cossegregação de um fenótipo com marcadores genéticos previamente mapeados.

Coativador. Proteína que coopera com um ativador de transcrição. No controle transcripcional eucariótico, os coativadores interagem fisicamente com ativadores transcripcionais e o aparelho basal de transcrição.

Código da histona. Modificação das histonas, como adição ou remoção dos grupos fosfato, grupos metila ou grupos acetila, que codificam informações afetando o modo como os genes são expressos.

Código genético degenerado. Refere-se ao fato de o código genético ter mais informação do que é necessário para especificar todos os 20 aminoácidos comuns.

Código genético não sobreposto. Refere-se ao fato de que, em geral, cada nucleotídio é uma parte de apenas um códon e codifica apenas um aminoácido em uma proteína.

Código genético universal. Refere-se ao fato de que determinados códons especificam os mesmos aminoácidos em quase todos os organismos.

Código de trincas. Refere-se ao fato de que três nucleotídios codificam cada aminoácido em uma proteína.

Codominância. Tipo de interação alélica na qual o heterozigoto expressa simultaneamente os traços de ambos homozigotos.

Códon. Sequência de três nucleotídios que codifica um aminoácido em uma proteína.

Códon com sentido. Códon que especifica um aminoácido em uma proteína.

Códon de iniciação. O códon no mRNA que especifica o primeiro aminoácido (fMet nas células bacterianas; Met nas células eucarióticas) de uma proteína; mais comumente AUG.

Códon de terminação. Códon no mRNA que sinaliza o final da tradução; também chamado de códon de parada. Os três códons de terminação comuns são UAA, UAG e UGA.

Códon sem sentido. Códon no mRNA que sinaliza o final da tradução; também chamado de códon de parada ou códon de terminação. Os 3 códons sem sentido mais comuns são UAA, UAG e UGA.

Códons sinônimos. Códons diferentes que especificam o mesmo aminoácido.

Coeficiente de coincidência. Razão de *crossing overs* duplos observados em relação aos *crossing overs* duplos esperados.

Coeficiente de correlação. Valor estatístico que mede o grau de associação entre duas ou mais variáveis. Um coeficiente de correlação pode variar de −1 a +1. Um valor positivo indica uma relação direta entre as variáveis; uma correlação negativa indica uma relação inversa. O valor absoluto do coeficiente de correlação fornece informações sobre a extensão da associação entre as variáveis.

Coeficiente de endogamia. Medida da endogamia; a probabilidade (variando de 0 a 1) de que dois alelos sejam idênticos por herança.

Coeficiente de regressão. Medida estatística que avalia como uma variável muda, em média, com uma mudança de unidade em outra variável.

Coeficiente de seleção. Medida da intensidade relativa de seleção contra um genótipo; igual a 1 menos a adaptação.

Coesina. Molécula que mantém as duas cromátides-irmãs de um cromossomo unidas. A degradação da coesina nos centrômeros permite que as cromátides se separem na anáfase da mitose e anáfase II da meiose.

Colapso do híbrido. Mecanismo de isolamento reprodutivo no qual espécies próximas são capazes de se acasalar e produzir descendentes F_1 viáveis e férteis, mas os genes não fluem entre as duas espécies, porque o cruzamento dos híbridos produz descendentes inviáveis ou estéreis.

Colinearidade. Conceito de que existe uma correspondência direta entre a sequência de nucleotídios de um gene e a sequência contínua de aminoácidos em uma proteína.

Colônia. Ver *Colônia bacteriana*.

Colônia bacteriana. Grupo de bactérias geneticamente idênticas derivadas de uma única célula bacteriana que sofre ciclos repetidos de divisão.

Compensação de dosagem. Equalização da quantidade de proteína produzida pelos genes ligados ao cromossomo X em

machos e fêmeas. Nos mamíferos placentários, a compensação de dosagem é possível graças à inativação aleatória de um cromossomo X nas células das fêmeas.

Complementação. Duas mutações diferentes na condição heterozigota são exibidas como o fenótipo do tipo selvagem; indica que as mutações estão em diferentes *loci*.

Complexo Antennapedia. Grupo de cinco genes homeóticos nas moscas-da-fruta que afeta o desenvolvimento da cabeça e dos segmentos torácicos anteriores da mosca adulta.

Complexo bitórax. Grupo de três genes homeóticos na *Drosophila* que influenciam os segmentos abdominal e torácico posterior na mosca adulta.

Complexo de iniciação 30S. Complexo inicial formado na iniciação da tradução nas células bacterianas; composto pela subunidade menor do ribossomo, mRNA, tRNA iniciador carreado com fMet, GTP e fatores de iniciação 1, 2 e 3.

Complexo de iniciação 70S. Complexo final formado na iniciação da tradução nas células bacterianas; consiste em subunidades pequenas e grandes do ribossomo, mRNA e tRNA iniciador carreado com fMet.

Complexo de ligação ao cap. Um grupo de proteínas nos eucariotos que se liga a 5′ cap e inicia a tradução. Auxilia na exportação do mRNA do núcleo para o citoplasma e promove o ciclo inicial (pioneiro) de tradução.

Complexo de remodelagem da cromatina. Complexo de proteínas que altera a estrutura da cromatina sem acetilação de histonas.

Complexo homeótico (HOM-C). Grupo principal de genes homeóticos nas moscas-da-fruta; composto pelo complexo Antennapedia, que afeta o desenvolvimento da cabeça e dos segmentos anteriores da mosca adulta, e o complexo bitórax, que afeta os segmentos torácico posterior e abdominal da mosca adulta.

Complexo silenciador induzido pelo RNA (RISC). Combinação de uma pequena molécula de RNA de interferência (siRNA) ou uma molécula de microRNA (miRNA) e proteínas que podem romper o mRNA, levando à sua degradação, ou afetando a transcrição, ou reprimindo a tradução do mRNA.

Complexo sinaptonêmico. Estrutura em três partes que se desenvolve entre cromossomos homólogos em sinapse.

Conceito biológico de espécie. Define uma espécie como um grupo de organismos cujos membros são capazes de cruzar entre si, mas são isolados do ponto de vista reprodutivo dos membros de outra espécie. Como diferentes espécies não trocam genes, cada espécie evolui de forma independente. Este conceito não é aceito por todos os biólogos.

Conceito de dominância. Princípio de hereditariedade descoberto por Mendel que afirma que quando dois alelos diferentes estão presentes em um genótipo, apenas um alelo pode ser expresso no fenótipo. O alelo dominante é o alelo que é expresso, e o alelo recessivo é o alelo que não é expresso.

Concordância. Porcentagem de pares de gêmeos nos quais ambos gêmeos têm um traço particular.

Concordante. Refere-se a um par de gêmeos em que ambos têm o traço em questão.

Configuração cis. Arranjo no qual dois ou mais genes do tipo selvagem estão em um cromossomo e seus alelos mutantes estão no cromossomo homólogo; também chamado de configuração de acoplamento.

Configuração de acoplamento. Ver *Configuração cis*.

Configuração trans. Arranjo no qual cada cromossomo tem um gene do tipo selvagem (dominante) e outro do tipo mutante (recessivo); também chamada de repulsão.

Conjugação. Mecanismo pelo qual o material genético pode ser trocado entre células bacterianas. Na conjugação, duas bactérias se aproximam, e forma-se uma conexão entre elas. Um plasmídio ou às vezes uma parte do cromossomo bacteriano passa por essa conexão de uma célula para outra.

Consanguinidade. Acasalamento entre indivíduos aparentados.

Contig. Conjunto de fragmentos de DNA sobrepostos, montado na ordem correta para formar uma extensão contínua da sequência de DNA.

Controle negativo. Regulagem de genes na qual a ligação de uma proteína reguladora ao DNA inibe a transcrição (a proteína reguladora é um repressor).

Controle positivo. Regulagem de genes na qual a ligação de uma proteína reguladora ao DNA estimula a transcrição (a proteína reguladora é um ativador).

Conversão gênica. Processo de troca genética não recíproca que pode produzir proporções anormais de gametas após a meiose.

Correpressor. Substância que inibe a transcrição em um sistema reprimível de regulação de gene, em geral uma pequena

molécula que se liga a uma proteína repressora e a altera de forma que o repressor é capaz de se ligar ao DNA e inibir a transcrição.

Corpúsculo de Barr. Estrutura condensada, de cor escura que é encontrada na maioria das células de mamíferos fêmeas placentárias e é um cromossomo X inativado.

Correlação fenotípica. Correlação entre dois ou mais fenótipos no mesmo indivíduo.

Correlação genética. Correlação fenotípica visto que os mesmos genes afetam duas ou mais características.

Correlação. Grau de associação entre duas ou mais variáveis.

Cosmídio. Vetor de clonagem que combina as propriedades dos plasmídios e vetores de fago e é usado para clonar pedaços grandes de DNA nas bactérias. Um cosmídio é um pequeno plasmídio que carreia um sítio λ *cos*, possibilitando que o plasmídio seja acondicionado em um revestimento viral.

Cotransdução. Processo no qual dois ou mais genes são transferidos juntos de uma célula bacteriana para outra. Apenas genes próximos no cromossomo bacteriano serão cotransduzidos.

Cotransformação. Processo no qual dois ou mais genes são transferidos juntos durante a transformação celular.

CRISPR RNAs (crRNAs). Pequenas moléculas de RNA encontradas nos procariotos que auxiliam na destruição de DNA estranho.

Cromátide acêntrica. Não tem centrômero; produzida quando o *crossing over* ocorre em uma inversão paracêntrica. A cromátide acêntrica não se liga à fibra do fuso e não se separa na meiose ou mitose; logo, ela é perdida após um ou mais ciclos de divisão celular.

Cromátide dicêntrica. A cromátide que tem dois centrômeros; produzida quando o *crossing over* ocorre em uma inversão paracêntrica. Os dois centrômeros da cromátide dicêntrica são puxados para polos opostos na mitose ou meiose, rompendo o cromossomo.

Cromátides-irmãs. Duas cópias de um cromossomo que são mantidas unidas no centrômero. Cada cromátide tem uma única molécula de DNA.

Cromatina. Material encontrado no núcleo eucariótico, composto por DNA e proteínas.

Cromossomo. Estrutura composta por DNA e proteínas associadas que carreia e transmite a informação genética.

Cromossomo acrocêntrico. Cromossomo no qual o centrômero está próximo da extremidade, produzindo um braço longo em uma extremidade e um botão ou satélite na outra extremidade.

Cromossomo artificial de levedura (YAC). Vetor de clonagem composto por uma molécula de DNA com uma origem de replicação, um par de telômeros e um centrômero de levedura. Os YACs carreiam segmentos muito grandes de DNA (em torno de muitas centenas de milhares de pares de base) e se replicam e separam como os cromossomos da levedura fazem.

Cromossomo bacteriano artificial (BAC). Vetor de clonagem usado nas bactérias capazes de carrear fragmentos de DNA do tamanho de 500 kb.

Cromossomo metacêntrico. Cromossomo no qual os dois braços têm aproximadamente o mesmo comprimento.

Cromossomo politênico. Cromossomo gigante nas glândulas salivares da *Drosophila melanogaster*. Cada cromossomo politênico consiste em várias moléculas de DNA lado a lado.

Cromossomo submetacêntrico. Cromossomo no qual o centrômero está deslocado para uma extremidade, criando um braço longo e um braço curto.

Cromossomo telocêntrico. Cromossomo no qual o centrômero está na extremidade ou muito perto dela.

Cromossomos sexuais. Cromossomos com diferença morfológica ou no número em machos e fêmeas.

Crossing over. Troca de material genético entre cromátides homólogas, mas não irmãs.

Crossing over desigual. Desalinhamento de duas moléculas de DNA durante o *crossing over*, resultando em uma molécula de DNA com uma inserção e outra com uma deleção.

Cruciforme. Estrutura formada pelo pareamento de repetições invertidas de ambas as fitas de DNA de fita dupla.

Cruzamento di-híbrido. Um cruzamento entre dois indivíduos com duas características diferentes – mais especificamente, um cruzamento entre indivíduos que são homozigotos para diferentes alelos em dois *loci* (*AA BB* × *aa bb*); também se refere a um cruzamento entre dois indivíduos que são heterozigotos em dois *loci* (*Aa Bb* × *Aa Bb*).

Cruzamento mono-híbrido. Um cruzamento entre dois indivíduos com uma única característica diferente – mais especificamente, um cruzamento entre indivíduos que são homozigotos para diferentes alelos no mesmo *locus* (*AA* × *aa*);

também se refere a um cruzamento entre dois indivíduos que são heterozigotos para dois alelos em um único *locus* (*Aa* × *Aa*).

Cruzamento-teste. Um cruzamento entre um indivíduo com genótipo desconhecido e um indivíduo com genótipo homozigoto recessivo.

Cruzamento-teste de dois pontos. Cruzamento entre um indivíduo heterozigoto em dois *loci* e um indivíduo homozigoto para os alelos recessivos nesses *loci*.

Cruzamento-teste de três pontos. Cruzamento entre um indivíduo heterozigoto em três *loci* e um indivíduo homozigoto para os alelos recessivos nesses *loci*.

Cruzamento tri-híbrido. Um cruzamento entre dois indivíduos com três características diferentes (*AA BB CC* × *aa bb cc*); também se refere a um cruzamento entre dois indivíduos heterozigotos em três *loci* (*Aa Bb Cc* × *Aa Bb Cc*).

Cruzamentos recíprocos. Cruzamentos nos quais os fenótipos dos genitores estão invertidos. Por exemplo, em um cruzamento, um macho alto é cruzado com uma fêmea baixa e, no outro, um macho baixo é cruzado com uma fêmea alta.

Decaimento de mRNA mediado por terminação sem sentido (NMD). Processo que leva à rápida eliminação do mRNA que tem um códon de parada prematuro.

Decaimento de RNA nonstop. Mecanismo nas células eucarióticas para lidar com ribossomos parados na extremidade 3′ de um mRNA que não tem um códon de terminação. Uma proteína se liga ao sítio A do ribossomo parado e recruta outras proteínas que degradam o mRNA a partir da extremidade 3′.

Decaimento no-go. Um sistema de vigilância do mRNA nos eucariotos que ajuda a remover ribossomos parados oriundos de estruturas secundárias no mRNA, dano químico ao mRNA, códons de parada prematura e defeitos nos ribossomos.

Deleção. Mutação na qual um ou mais nucleotídios são retirados de uma sequência de DNA.

Deleção cromossômica. Perda de um segmento de cromossomo.

Deleção no quadro de leitura. Deleção de alguns múltiplos de três nucleotídios, que não altera o quadro de leitura do gene.

Depressão da endogamia. Adaptabilidade reduzida que surge a partir da procriação consanguínea; em muitos casos por causa da expressão aumentada de traços recessivos letais ou deletérios.

Depurinação. Quebra da ligação covalente que conecta uma base purina ao carbono 1′ da desoxirribose, resultando na perda da base purina.

Deriva genética. Mudança na frequência alélica por causa de um erro de amostragem.

Desaminação. Perda de um grupo amino (NH_2) de uma base.

Descendentes não recombinantes (parentais). Têm as combinações originais dos traços presentes nos genitores.

Descendentes parentais. Ver *Descendentes não recombinantes (parentais).*

Descendentes recombinantes. Descendentes com novas combinações de traços formados a partir de gametas recombinantes.

Desequilíbrio de ligação. Associação não aleatória entre as variantes genéticas dentro de um haplótipo.

Deserto gênico. Em referência à densidade de genes no genoma, uma região pobre em genes – ou seja, um longo segmento de DNA que consiste em centenas de milhares a milhões de pares de bases completamente isentos de qualquer gene ou outras sequências funcionais.

Deslizamento da fita. Deslize do molde e das fitas recém-sintetizadas na replicação no qual uma das fitas se solta da outra e os nucleotídios são inseridos ou apagados na fita recém-sintetizada.

Desnaturação. Processo que separa as duas fitas do DNA quando ele é aquecido.

Desoxirribonucleotídio. Elemento básico do DNA, composto por desoxirribose, um fosfato e uma base nitrogenada.

Desoxirribose. Açúcar de cinco carbonos no DNA; não tem o grupo hidroxila no átomo de carbono 2′.

Desvio padrão. Medida estatística que descreve a variabilidade de um grupo de medidas; a raiz quadrada da variância.

Determinação. Processo pelo qual uma célula se torna comprometida no desenvolvimento de um tipo celular específico.

Determinação do sexo. Especificação do sexo (macho ou fêmea). Mecanismos para determinar o sexo incluem sistemas cromossômicos, gênicos e ambientais.

Determinação do sexo pelo gene. A determinação do sexo na qual o fenótipo sexual é especificado por genes em um ou mais *loci*, mas não existem diferenças claras nos cromossomos dos machos e das fêmeas.

Diacinese. Quinto estágio da prófase I na meiose. Na diacinese, os cromossomos se contraem, a membrana nuclear se rompe e o

fuso se forma.

Diagnóstico genético pré-implante (PGD). Exame genético em um embrião produzido pela fertilização *in vitro* antes de implantá-lo no útero.

Diagnóstico genético pré-natal não invasivo. Um teste genético realizado em um feto antes do nascimento sem tirar uma amostra de tecido do feto. Geralmente feito ao testar o DNA fetal encontrado no sangue materno.

Difração por raios X. Método para analisar o formato tridimensional e a estrutura das substâncias químicas. Cristais de uma substância são bombardeados com raios X, que batem nos cristais, ricocheteiam e produzem um padrão de difração no detector. O padrão dos pontos produzidos no detector fornece informações sobre a estrutura da molécula.

Dímero de pirimidina. Estrutura na qual uma ligação se forma entre duas moléculas adjacentes de pirimidina na mesma fita de DNA; rompe a ponte de hidrogênio normal entre as bases complementares e distorce a configuração normal da molécula de DNA.

Diploide. Tem dois conjuntos de cromossomos (dois genomas)

Diploide parcial. Célula bacteriana que tem duas cópias de genes, incluindo uma cópia de cromossomo bacteriano e a outra em um pedaço extra de DNA (em geral um plasmídio); também chamado de merozigoto.

Diplóteno. Quarto estágio da prófase I na meiose. No diplóteno, os centrômeros dos cromossomos homólogos se distanciam, mas os homólogos permanecem presos nos quiasmas.

Discordante. Refere-se a um par de gêmeos dos quais um tem o traço em questão e o outro não.

Disgenesia híbrida. Surgimento súbito de várias mutações, aberrações cromossômicas e esterilidade nos descendentes de um cruzamento entre uma mosca macho que tem elementos *P* e uma fêmea que não tem esses elementos.

Dissomia uniparental. Herança de ambos os cromossomos de um par homólogo de um único genitor.

Distribuição de frequência. Gráfico dos valores representados. Na genética, os fenótipos encontrados em um grupo de indivíduos são deslocados como uma distribuição de frequência. Tipicamente, os fenótipos são inseridos no eixo horizontal (x) e os números (ou proporções) de indivíduos com cada fenótipo são inseridos no eixo vertical (y).

Distribuição independente. Separação independente dos pares de cromossomos na anáfase I da meiose; contribui para a variação genética.

Distribuição normal. Tipo comum de distribuição de frequência que exibe uma curva simétrica, em forma de sino; em geral, surge quando vários fatores independentes contribuem para a medida.

Diversidade juncional. Adição ou deleção de nucleotídios nas junções dos segmentos de gene unidos na recombinação somática dos genes que codificam anticorpos e receptores de linfócito T.

DNA altamente repetitivo. DNA composto por sequências curtas, encontradas em centenas de milhares a milhões de cópias; agrupadas em algumas regiões dos cromossomos.

DNA de cloroplasto (cpDNA). DNA nos cloroplastos; tem muitas características em comum com o DNA eubacteriano e consiste tipicamente em uma molécula circular que não tem as proteínas histona e codifica alguns rRNAs, tRNAs e proteínas encontradas nos cloroplastos.

DNA de sequência única. Sequência presente apenas uma vez ou algumas poucas vezes em um genoma.

DNA girase. Enzima topoisomerase da *E. coli* que alivia a fita de torção que se constrói à frente da forquilha de replicação.

DNA helicase. Enzima que abre o DNA de fita dupla ao romper as pontes de hidrogênio.

DNA heteroduplex. DNA com duas fitas, cada uma oriunda de um cromossomo diferente.

DNA ligador. Trecho do DNA que separa dois nucleossomos.

DNA ligase. Enzima que catalisa a formação de uma ligação fosfodiéster entre os grupos 3′-OH e 5′-fosfato adjacentes em uma molécula de DNA.

DNA mitocondrial (mtDNA). DNA presente nas mitocôndrias; tem muitas características em comum com o DNA eubacteriano e consiste tipicamente em uma molécula circular que não tem as proteínas histona e codifica alguns rRNAs, tRNAs e proteínas encontradas nas mitocôndrias.

DNA moderadamente repetitivo. DNA composto por sequências de 150 a 300 pb de comprimento e que são repetidas milhares de vezes.

DNA polimerase. Enzima que sintetiza o DNA.

DNA polimerase α. DNA polimerase eucariótica que inicia a replicação.

DNA polimerase β. DNA polimerase eucariótica que participa do reparo ao DNA.

DNA polimerase γ. DNA polimerase eucariótica que replica o DNA mitocondrial. Uma DNA polimerase semelhante a γ replica o DNA do cloroplasto.

DNA polimerase δ. DNA polimerase eucariótica que replica a fita de revestimento durante a síntese de DNA; também faz o reparo do DNA e a síntese translesão do DNA.

DNA polimerase ε. DNA polimerase eucariótica que replica a fita líder durante a síntese de DNA.

DNA polimerase I. DNA polimerase bacteriana que remove e substitui os *primers* de RNA com nucleotídios de DNA.

DNA polimerase II. DNA polimerase bacteriana que participa do reparo do DNA; reinicia a replicação após a síntese ter parado por causa do dano ao DNA.

DNA polimerase III. DNA polimerase bacteriana que sintetiza novas cadeias de nucleotídios ao usar o grupo 3′-OH fornecido pelo *primer*.

DNA polimerase IV. DNA polimerase bacteriana; é provável que participe do reparo ao DNA.

DNA polimerase V. DNA polimerase bacteriana; é provável que participe do reparo ao DNA.

DNA polimerase translesão. DNA polimerases especializadas que são capazes de se replicar por meio de estruturas distorcidas e lesões volumosas que prejudicam outras DNA polimerases. Em geral elas produzem mais erros durante a síntese de DNA do que outras DNA polimerases.

DNA repetitivo. Sequências que existem em cópias múltiplas em um genoma.

Doença autoimune. Caracterizada pela resposta imunológica anormal ao próprio antígeno da pessoa.

Dogma central. Conceito de que a informação genética passa do DNA para o RNA para a proteína em uma via de informação de sentido único.

Dominância completa. Refere-se a um alelo ou um fenótipo que é expresso em homozigotos (*AA*) e heterozigotos (*Aa*); apenas o alelo dominante é expresso em um fenótipo heterozigoto.

Dominância incompleta. Refere-se ao fenótipo de um heterozigoto que é intermediário entre os fenótipos dos dois homozigotos.

Dominante. Refere-se a um alelo ou um fenótipo que é expresso em homozigotos (*AA*) e heterozigotos (*Aa*); apenas o alelo dominante é expresso em um fenótipo heterozigoto.

Domínio. Parte funcional de uma proteína.

Domínio da proteína. Região de uma proteína que tem formato ou função específicos.

Domínio de ativação rico em aminoácido ácido. Comumente encontrado em algumas proteínas ativadoras de transcrição, um domínio com múltiplos aminoácidos com cargas negativas e estimula a transcrição de alguns genes.

Duplicação cromossômica. Mutação que duplica um segmento de um cromossomo.

Duplicação deslocada. Rearranjo cromossômico no qual o segmento duplicado está distante do segmento original, tanto no mesmo cromossomo ou em um cromossomo diferente.

Duplicação em tandem. Rearranjo cromossômico no qual um segento de cromossomo duplicado está adjacente ao segmento original.

Duplicação reversa. Duplicação de um segmento de cromossomo no qual a sequência do segmento duplicado é invertida em relação à sequência do segmento original.

Duplicações segmentares. Regiões maiores que 1.000 pb com sequências quase idênticas nos genomas eucarióticos.

Edição do RNA. Processo no qual a sequência codificadora de proteína de um mRNA é alterada após a transcrição. Os aminoácidos especificados pelo mRNA alterado são diferentes dos previstos a partir da cadeia de nucleotídios do gene que codifica a proteína.

Efeito da posição. Dependência da expressão de um gene sobre a localização do gene no genoma.

Efeito do fundador. Erro de amostragem que surge quando uma população é estabelecida por um pequeno número de indivíduos; leva à deriva genética.

Efeito genético materno. Determina o fenótipo de um descendente. Com o efeito genético materno, um descendente herda os genes para as características de ambos os genitores, mas o fenótipo do descendente é determinado pelo genótipo nuclear de sua

mãe, e não por seu próprio genótipo.

Elemento autônomo. Elemento de transposição que está totalmente funcional e é capaz de se transpor.

Elemento curto intercalado (SINE). Sequência curta de DNA repetida várias vezes e intercalada em todo o genoma.

Elemento de reconhecimento TFIIB (BRE). Sequência consenso [(G ou C)(G ou C)(G ou C)CGCC] encontrada no cerne de alguns promotores de RNA polimerase II; em geral localizada a partir de 32 a 38 pb *upstream* do sítio de início de transcrição.

Elemento de resposta. Sequência de DNA comum encontrada antecedente (*upstream*) a alguns grupos de genes eucarióticos. Uma proteína reguladora se liga a um elemento de resposta e estimula a transcrição de um gene. A existência do mesmo elemento de resposta em vários promotores ou acentuadores possibilita que um único fator estimule simultaneamente a transcrição de vários genes.

Elemento de transposição. Sequência de DNA capaz de se deslocar de um sítio para outro no genoma por meio de um mecanismo diferente da recombinação homóloga.

Elemento posterior (downstream) ao promotor central. Sequência consenso [RG(A ou T)CGTG] encontrada no cerne de alguns promotores da RNA polimerase II eucariótica; em geral localizada aproximadamente 30 pb *downstream* do sítio de início da transcrição.

Elemento longo intercalado (LINE). Sequência longa de DNA repetida várias vezes e dispersa por todo o genoma.

Elemento não autônomo. Elemento de transposição que não pode se transpor mas pode transpor na presença do elemento de transposição autônomo da mesma família.

Elemento regulador. Sequência de DNA que afeta a transcrição de outras sequências de DNA às quais está fisicamente ligada.

Elemento upstream. Sequência consenso encontrada em alguns promotores bacterianos com vários pares A-T e encontrado cerca de 40 a 60 pb anterior ou *upstream* ao sítio de início da transcrição.

Eletroforese bidimensional em gel de poliacrilamida (2D-PAGE). Método para separar proteínas em pontos nos quais as proteínas são separadas em uma dimensão pela carga, separadas em uma segunda dimensão pela massa e então coradas. Cada ponto é proporcional à quantidade de proteína presente.

Eletroforese em gel. Técnica usada para separar moléculas carreadas (como proteínas ou ácidos nucleicos) com base no tamanho molecular ou carga ou ambos.

Embaralhamento de éxons. Processo importante na evolução dos genes eucarióticos, no qual éxons de diferentes genes são trocados e misturados em novas combinações, criando novos genes que são mosaicos de outros genes preexistentes.

Endonuclease de restrição. Termo técnico para uma enzima de restrição que identifica sequências de base particulares no DNA e faz cortes de fita dupla próximos.

Endogamia. Acasalamento entre indivíduos consanguíneos que ocorre mais frequentemente do que esperado com base no acaso.

Engenharia genética. Termo comum para a tecnologia de DNA recombinante.

Envelope nuclear. Membrana que circunda o material genético nas células eucarióticas para formar um núcleo; separa o DNA do restante do conteúdo celular.

Enzima de restrição. Identifica sequências de bases particulares no DNA e faz cortes de fita dupla próximos; também chamada de endonuclease de restrição.

Enzima modificadora de tRNA. Cria uma base modificada no RNA ao catalisar uma mudança química na base padrão.

Epialelos. Alelos sem diferença na sua sequência de bases, mas com diferenças epigenéticas que produzem variações hereditárias nos fenótipos.

Epigenética. Fenômenos causados por alterações no DNA que não incluem mudanças na sequência de bases; em geral, afetam a maneira na qual as sequências de DNA são expressas. Tais mudanças são estáveis e podem ser herdadas de modo que são transmitidas de uma célula para outra.

Epigenoma. Todas as modificações epigenéticas dentro do genoma de um organismo.

Episssomo. Plasmídio capaz de se integrar ao cromossomo bacteriano.

Epistasia. Tipo de interação genética na qual um gene em um *locus* mascara ou suprime os efeitos de um gene em um *locus* diferente.

Equilíbrio. Situação na qual não ocorre mais mudança; na genética populacional, refere-se a uma população na qual as frequências alélicas não se alteram.

Equilíbrio de Hardy-Weinberg. Frequências de genótipos quando as condições da lei de Hardy-Weinberg são satisfeitas.

Erro de amostragem. Desvios das proporções esperadas por causa de ocorrências ao acaso quando o número de eventos é pequeno.

Erro de incorporação. Incorporação de um nucleotídio danificado ou par de bases não pareado em uma molécula de DNA.

Erro de replicação. Replicação de um erro de incorporação no qual uma mudança na sequência de DNA foi replicada e todas as bases que pareiam na nova molécula de DNA estão corretas.

Especiação. Processo pelo qual uma nova espécie surge. (Ver também *Conceito biológico de espécie, Especiação alopátrica* e *Especiação simpátrica*.)

Especiação alopátrica. Especiação que surge quando uma barreira geográfica divide uma população em dois grupos e bloqueia a troca de genes entre eles.

Especiação simpátrica. Especiação que surge na ausência de qualquer barreira geográfica para o fluxo dos genes; mecanismos de isolamento reprodutivo evoluem dentro de uma única população que se intercruza.

Espécie. Termo aplicado a diferentes gêneros ou tipos de organismos vivos.

Espectrometria de massa. Método para determinar com precisão a massa molecular de uma molécula ao usar a taxa de migração de uma molécula ionizada em um campo elétrico.

Espermátide. Produto imediato da meiose II na espermatogênese; amadurece em espermatozoide.

Espermatócito primário. Espermatogônia que entrou na prófase I.

Espermatócito secundário. Produto da meiose I em animais machos.

Espermatogênese. Produção de espermatozoides em animais.

Espermatogônia. Célula diploide no testículo; capaz de sofrer meiose para produzir um espermatozoide.

Esporófito. Fase diploide do ciclo de vida nas plantas.

Estado relaxado do DNA. Estado de energia de uma molécula de DNA quando não há tensão estrutural na molécula.

Esterilidade híbrida. Embriões híbridos completam o desenvolvimento, mas são estéreis; exemplificados pelo acasalamento entre burros e cavalos para produzir uma mula, um descendente viável mas geralmente estéril.

Estrutura de folha de trevo. Estrutura secundária comum a todos tRNAs.

Estrutura lariat. Estrutura semelhante à alça criada no *splicing* de pré-mRNA nuclear no qual a extremidade 5′ de um íntron é fixada a um ponto de ramificado no pré-mRNA.

Estrutura primária de uma proteína. A sequência de aminoácidos de uma proteína.

Estrutura quaternária de uma proteína. Interação de dois ou mais polipeptídios para formar uma proteína funcional.

Estrutura secundária de uma proteína. Arranjo de dobra regular de aminoácidos em uma proteína. Estruturas secundárias comuns encontradas nas proteínas incluem a alfa-hélice e a folha beta dobrada.

Estrutura terciária de uma proteína. Dobra de ordem superior de aminoácidos em uma proteína para criar a forma tridimensional da molécula.

Estudos de associação genômica ampla. Estudos que procuram por associações não aleatórias entre a presença de um traço e alelos em muitos *loci* diferentes dispersos em um genoma – ou seja, para associações entre traços e localizações particulares dos alelos em uma população.

Etiqueta de sequência expressa (EST). Fragmento singular de DNA proveniente da região codificadora de um gene, produzido pela transcrição reversa do RNA celular. Partes do fragmento são sequenciadas de modo que possam ser identificadas.

Eubactérias. Uma das três divisões principais da vida. As eubactérias são organismos unicelulares com células procarióticas e incluem a maioria das bactérias comuns.

Eucarioto. Organismo com uma estrutura celular complexa, incluindo um envelope nuclear e organelas ligadas à membrana. Uma das três principais divisões da vida, os eucariotos incluem formas unicelulares e multicelulares.

Eucromatina. Cromatina que sofre condensação e descondensação no curso do ciclo celular.

Evolução. Mudança genética que ocorre em um grupo de organismos.

Evolução clonal. Processo pelo qual as mutações que potencializam a capacidade das células em se proliferar predominam em um clone de células, possibilitando que o clone se torne muito rápido no crescimento e mais agressivo nas propriedades

proliferativas.

Exame genético pré-sintomático. Testa as pessoas para determinar se elas herdaram um gene responsável por uma doença antes que os sintomas da doença apareçam.

Exocruzamento (*outcrossing*). Acasalamento entre indivíduos não consanguíneos que ocorre mais frequentemente do que esperado com base no acaso.

Éxon. Região codificadora de um gene partido (um gene que é interrompido por íntrons). Após o processamento, os éxons permanecem no RNA mensageiro.

Expressividade. Grau no qual um traço é expresso.

Extremidade 3′. Extremidade de uma cadeia de polinucleotídios onde um grupo OH é preso ao átomo do carbono 3′ do nucleotídio.

Extremidade 5′. Extremidade da cadeia de polinucleotídios onde um fosfato é preso ao átomo 5′ carbono do nucleotídio.

Extremidade coesiva. Extremidade protuberante, curta, de fita única em uma molécula produzida quando o DNA é cortado por algumas enzimas de restrição. As extremidades coesivas são complementares e podem parear espontaneamente para religar fragmentos de DNA que foram cortados pela mesma enzima de restrição.

Fago. Ver *Bacteriófago*.

Fago temperado. Bacteriófago que utiliza o ciclo lisogênico, no qual o DNA do fago se integra ao cromossomo bacteriano e permanece no estado inativo.

Fago transdutor. Tem um pedaço de cromossomo bacteriano dentro do revestimento do fago. (Ver também *Transdução generalizada*.)

Fago virulento. Bacteriófago que se reproduz apenas por meio do ciclo lítico e destrói sua célula hospedeira.

Família de genes. Ver *Família de multigenes.*

Família de multigenes. Conjunto de genes com sequência semelhante que surgem a partir de eventos de duplicação repetidos; muitas vezes codificam diferentes proteínas.

Fase M (mitótica). Período de divisão celular ativa; inclui mitose (divisão nuclear) e citocinese (divisão citoplasmática).

Fase S (síntese). Estágio da interfase no ciclo celular. Na fase S, o DNA é duplicado.

Fator de iniciação 1 (IF-1). Proteína necessária para a iniciação da tradução nas células bacterianas; potencializa a dissociação das subunidades maiores e menores do ribossomo.

Fator de iniciação 2 (IF-2). Proteína necessária para a iniciação da tradução nas células bacterianas; forma um complexo com GTP e o tRNA iniciador carreado e então libera o tRNA carreado para o complexo de iniciação.

Fator de iniciação 3 (IF-3). Proteína necessária para o início da tradução nas células bacterianas; liga-se à subunidade menor do ribossomo e evita que a subunidade maior se ligue durante a iniciação.

Fator de liberação. Proteína necessária para terminar a tradução; liga-se a um ribossomo quando o códon de parada é alcançado e estimula a liberação da cadeia de polipeptídios, o tRNA e o mRNA do ribossomo. As células eucarióticas precisam de dois fatores de liberação (eRF-1 e eRF-2), enquanto a *E. coli* precisa de três (RF-1, RF-2 e RF-3).

Fator de licenciamento para replicação. Proteína que garante que a replicação ocorre apenas uma vez em cada origem; necessária na origem antes que a replicação possa ser iniciada e removida após o DNA ser replicado.

Fator de transcrição. Proteína que se liga a sequências de DNA nas células eucarióticas e afeta a transcrição.

Fator de transcrição geral. Proteína que se liga a promotores eucarióticos próximos ao sítio de início da transcrição e é uma parte do aparelho basal de transcrição que inicia a transcrição.

Fator F (fertilidade). Episssomo da *E. coli* que controla a conjugação e troca de genes entre células de *E. coli*. O fator F tem uma origem de replicação e genes que permitem que a bactéria sofra conjugação.

Fator G de alongamento (EF-G). Proteína que se combina com GTP e é necessária para o movimento do ribossomo ao longo do mRNA durante a tradução.

Fator promotor da mitose (MPF). Proteína que atua no controle do ciclo celular; consiste em uma ciclina combinada com quinase dependente de ciclina (CDK). O MPF ativo estimula a mitose.

Fator rho. Subunidade da RNA polimerase bacteriana que facilita o término da transcrição de alguns genes.

Fator sigma. Subunidade da RNA polimerase bacteriana que permite que a RNA polimerase identifique um promotor e inicie a

transcrição.

Fator Ts de alongamento (EF-Ts). Proteína que regenera o fator Tu de alongamento no estágio de alongamento da síntese proteica.

Fator Tu de alongamento (EF-Tu.) Proteína que faz parte do estágio de alongamento da síntese proteica, forma um complexo com GTP e um aminoácido carreado e então libera o tRNA carreado para o ribossomo.

Fenilcetonúria (PKU). Doença genética caracterizada por déficit intelectual, pele clara e eczema; causada por mutações no gene que codifica a fenilalanina hidroxilase (PAH), uma enzima hepática que normalmente metaboliza o aminoácido fenilalanina. Quando a enzima é defeituosa, a fenilalanina não é metabolizada e se acumula no corpo, consequentemente levando a déficit intelectual e outras características da doença. É herdada como uma doença autossômica recessiva e pode ser tratada de forma eficiente ao limitar a fenilalanina na dieta.

Fenocópia. Fenótipo que é produzido por efeitos ambientais e é o mesmo que o fenótipo produzido por um genótipo.

Fenótipo. Surgimento ou manifestação de uma característica.

Fertilização. Fusão de gametas (células sexuais) para formar um zigoto.

Fertilização dupla. Fertilização em plantas; inclui a fusão de uma célula espermática com um óvulo para formar um zigoto e a fusão da segunda célula espermática com os núcleos polares para formar um endosperma.

Filogenia. Relações evolutivas em um grupo de organismos ou genes, em geral descritas como uma árvore genealógica ou diagrama de ramificação.

Fingerprinting do DNA. Técnica usada para identificar indivíduos ao examinar suas sequências de DNA.

Fita líder. Fita de DNA que é replicada de forma contínua.

Fita molde. A fita de DNA usada como molde durante a transcrição. O RNA sintetizado durante a transcrição é complementar e antiparalelo à fita molde.

Fita não molde. A fita de DNA que é complementar à fita molde; não é usada, de forma geral, como molde durante a transcrição.

Fita tardia. Fita de DNA que é replicada de forma descontínua.

Fitas complementares do DNA. A relação entre as duas fitas de nucleotídios do DNA na qual cada purina em uma fita pareia com uma pirimidina específica da fita oposta (A pareia com T, e G pareia com C).

Fixação. Ponto no qual um alelo alcança a frequência 1. Nesse ponto, todos os membros da população são homozigotos para o mesmo alelo.

Fluxo gênico. Movimento de genes de uma população para outra; também chamado de migração.

Forquilha de replicação. Ponto no qual uma molécula de DNA de fita dupla se separa em duas fitas únicas que servem como moldes para a replicação.

Fosfodiéster. Molécula com R-O-P-O-R, em que R é o grupo contendo carbono, O é oxigênio e P é fósforo.

Fragmento de Okazaki. Trechos curtos de DNA recém-sintetizado. Produzidos pela replicação descontínua na fita tardia, esses fragmentos são, consequentemente, unidos.

Frequência alélica. Proporção de um alelo particular em uma população.

Frequência de recombinação. Proporção de descendentes recombinantes produzidos em um cruzamento.

Frequência genotípica. Proporção de um genótipo particular em uma população.

Função de mapeamento. Relaciona as frequências de recombinação com as reais distâncias físicas entre os genes.

Fuso mitótico. Arranjo de microtúbulos que se irradiam a partir de dois polos; move os cromossomos na mitose e na meiose.

G_0 (gap 0). Estágio sem divisão do ciclo celular.

G_1 (gap 1). Estágio na interfase do ciclo celular no qual a célula cresce e se desenvolve.

G_2 (gap 2). Estágio da interfase no ciclo celular que vem após a replicação do DNA. Em G_2, a célula se prepara para a divisão.

Gameta não balanceado. Gameta com número variável de cromossomos; alguns cromossomos podem estar ausentes e outros podem estar presentes em mais de uma cópia.

Gameta não recombinante (parental). Tem apenas as combinações originais dos genes presentes nos genitores.

Gameta parental. Ver *Gameta não recombinante* (*parental*).

Gameta recombinante. Gameta com novas combinações de genes.

Gametófito. Fase haploide do ciclo de vida nas plantas.

Gargalo genético. Erro de amostragem que surge quando o tamanho de uma população é reduzido drasticamente; leva à deriva genética.

Gêmeos dizigóticos. Gêmeos não idênticos que surgem quando dois óvulos diferentes são fertilizados por dois espermatozoides diferentes; também são chamados de gêmeos fraternos.

Gêmeos fraternos. Gêmeos não idênticos que surgem quando dois óvulos diferentes são fertilizados por dois espermatozoides diferentes; também são chamados de gêmeos dizigóticos.

Gêmeos idênticos. Gêmeos que surgem quando um único óvulo fertilizado por um único espermatozoide se divide em dois embriões separados; também chamados de gêmeos monozigóticos.

Gêmeos monozigóticos. Gêmeos idênticos que surgem quando um único óvulo fertilizado por um único espermatozoide se divide em dois embriões separados.

Gêmeos não idênticos. Gêmeos que surgem quando dois óvulos diferentes são fertilizados por dois espermatozoides diferentes; também são chamados de gêmeos dizigóticos ou fraternos.

Gene. Fator genético que ajuda a determinar um traço; definido em nível molecular como uma sequência de DNA que é transcrita em uma molécula de RNA.

Gene constitutivo. Um gene que não é regulado e é expresso de forma contínua.

Gene da polaridade do ovo. Determina os principais eixos do desenvolvimento em um embrião inicial da mosca-da-fruta. Um conjunto de genes da polaridade do ovo determina o eixo anteroposterior, e outro conjunto determina o eixo dorsoventral.

Gene da região determinante do sexo no cromossomo Y (*SRY*). No cromossomo Y, um gene que dispara o desenvolvimento masculino.

Gene epistático. Mascara ou suprime o efeito de um gene em um *locus* diferente.

Gene estrutural. Sequência de DNA que codifica uma proteína atuante no metabolismo ou na biossíntese ou que tem um papel estrutural na célula.

Gene haploinsuficiente. Deve estar presente em duas cópias para a função normal. Se faltar uma das cópias do gene, é produzido um fenótipo mutante.

Gene hipostático. Gene que é mascarado ou suprimido pela ação de um gene em um *locus* diferente.

Gene homeótico. Determina a identidade de segmentos individuais ou partes no embrião inicial. As mutações em tais genes fazem com que partes do corpo surjam nos locais errados.

Gene Hox. Gene que tem um *homeobox*.

Gene regulador. Gene associado a um óperon em células bacterianas que codifica uma proteína ou molécula de RNA que controla a transcrição de um ou mais genes estruturais.

Gene regulatório. Sequência de DNA que codifica uma proteína ou molécula de RNA que interage com sequências de DNA e afeta sua transcrição ou tradução ou ambos.

Gene supressor de tumor. Gene que normalmente inibe a divisão celular. Mutações recessivas nesses genes contribuem para o câncer.

Genes de polaridade segmentar. Conjunto de genes de segmentação nas moscas-da-fruta que afetam a organização dos segmentos. Mutações nesses genes fazem com que parte de cada segmento seja apagado e substituído por uma imagem-espelho de parte ou todo de um segmento adjacente.

Genes de segmentação. Conjunto de cerca de 25 genes nas moscas-da-fruta que controlam a diferenciação do embrião em segmentos individuais, afetando o número e a organização dos segmentos. Mutações nesses genes, em geral, eliminam grupos inteiros de segmentos.

Genes da regra dos pares. Conjunto de genes de segmentação nas moscas-da-fruta que define as seções regionais do embrião e afeta os segmentos alternados. Mutações nesses genes levam à deleção de todos os outros segmentos.

Genes gap. Nas moscas-da-fruta, conjunto de genes de segmentação que definem grandes seções do embrião. Mutações nesses genes, em geral, eliminam grupos inteiros de segmentos adjacentes.

Genes homólogos. Genes próximos do ponto de vista evolutivo, que descenderam de um gene de um ancestral comum.

Genes ligados. Genes localizados no mesmo cromossomo.

Genes ortólogos. Genes homólogos encontrados em diferentes espécies porque as duas espécies têm um ancestral em comum que também tem o gene.

Genes parálogos. Genes homólogos na mesma espécie que surgem por meio da duplicação de um único gene ancestral.

Genética da transmissão. Campo da genética que inclui os princípios básicos da genética e como os traços são herdados.

Genética direta. Abordagem tradicional para o estudo da função do gene que começa com um fenótipo (um organismo mutante) e prossegue para um gene que codifica o fenótipo.

Genética molecular. Estudo da natureza química da informação genética e como é codificada, replicada e expressa.

Genética populacional. Estudo da composição genética das populações (grupos de membros da mesma espécie) e como um grupo coletivo de genes de uma população muda com o tempo.

Genética quantitativa. Análise genética de características complexas ou características influenciadas por múltiplos fatores genéticos.

Genética reversa. Uma abordagem molecular que inicia com um genótipo (uma sequência de DNA) e continua com o fenótipo ao alterar a sequência ou inibir sua expressão.

Genoma. Conjunto completo de instruções genéticas para um organismo.

Genômica. Estudo do conteúdo, da organização e da função da informação genética nos genomas inteiros.

Genômica comparativa. Estudos comparativos dos genomas de diferentes organismos.

Genômica estrutural. Área da genômica que estuda a organização e a sequência de informações existentes nos genomas; às vezes, é usada por químicos de proteínas para se referir à determinação da estrutura tridimensional das proteínas.

Genômica funcional. Área da genômica que estuda as funções da informação genética contida nos genomas.

Genótipo. O conjunto de genes existentes em um organismo.

Geração F_1 (primeira geração filial). Descendentes dos progenitores iniciais (P) em um cruzamento genético.

Geração F_2 (segunda geração filial). Descendentes da geração F_1 em um cruzamento genético; a terceira geração de um cruzamento genético.

Geração P (parental). Primeiro conjunto de genitores em um cruzamento genético.

Ginandromorfia. Organismo que é um mosaico para os cromossomos sexuais, tendo tecidos com diferentes constituições de cromossomos sexuais.

Girase. Ver *DNA girase*.

Grampo. Do inglês, *hairpin*. Estrutura secundária formada quando cadeias de nucleotídios na mesma fita são complementares e pareiam entre si.

Grupo de ligação. Genes localizados juntos no mesmo cromossomo.

Grupo fosfato. Um átomo de fósforo ligado a quatro átomos de oxigênio; um dos três componentes de um nucleotídio.

Guanina (G). Base purina no DNA e do RNA.

Haploide. Organismo com um único conjunto de cromossomos (um genoma).

Haploinsuficiência. O surgimento de um fenótipo mutante em uma célula ou organismo que é heterozigoto para um traço normalmente recessivo.

Haplótipo. Um conjunto específico de variantes genéticas ou alelos ligados em um único cromossomo ou em parte de um cromossomo.

H-DNA. Estrutura de DNA com três cadeias de nucleotídios (DNA triplex). Pode ocorrer quando uma única cadeia de nucleotídios de uma parte da molécula pareia com o DNA de fita dupla de outra parte da molécula.

Helicase. Ver *DNA helicase*.

Hemizigosidade. Tem um único alelo em um *locus*. Machos dos organismos com determinação sexual XX-XY são hemizigotos para os *loci* ligados ao X porque suas células têm um único cromossomo X.

Herança citoplasmática. Herança das características codificadas pelos genes localizados no citoplasma. Como em geral o citoplasma é totalmente fornecido por apenas um genitor, a maioria das características herdadas do citoplasma é herdada de um único genitor.

Herança das características adquiridas. Noção inicial de herança que propõe que os traços adquiridos são transmitidos para

os descendentes.

Herança por mistura. Antigo conceito de hereditariedade que propõe que os descendentes têm uma mistura de traços de ambos os genitores.

Herdabilidade. Proporção da variação fenotípica devido a diferenças genéticas. (Ver também *Herdabilidade de sentido amplo* e *Herdabilidade de sentido restrito*.)

Herdabilidade de sentido amplo. Proporção da variação fenotípica que pode ser atribuída à variância genética.

Herdabilidade de sentido restrito. Proporção de variância fenotípica que pode ser atribuída à variância genética aditiva.

Herdabilidade realizada. Herdabilidade de sentido restrito medida a partir de um experimento de resposta à seleção.

Heredograma. Representação por esquema do histórico familiar destacando a herança de um ou mais traços ou doenças.

Hermafroditismo. Condição na qual um organismo tem as estruturas reprodutoras masculina e feminina. Os hermafroditas verdadeiros produzem os gametas masculino e feminino.

Hermafroditismo sequencial. Fenômeno no qual o sexo de um organismo muda no curso de sua vida; o organismo é macho em um período ou estágio de desenvolvimento e fêmea em um diferente período ou estágio.

Heterocárion. Célula com dois núcleos derivados de diferentes células por meio da fusão celular.

Heterocromatina. Cromatina que permanece no estado altamente condensado durante todo o ciclo celular; encontrada nos centrômeros e telômeros da maioria dos cromossomos.

Heteroplasmia. Presença de duas ou mais variantes diferentes de DNA no citoplasma de uma única célula.

Heterozigose. Refere-se a um organismo que tem dois alelos diferentes em um *locus*

Heterozigoto composto. Indivíduo com dois alelos recessivos diferentes em um *locus* que resulta em um fenótipo recessivo.

Hibridização. Pareamento de duas cadeias de nucleotídios de fita única parcial ou totalmente complementares.

Hibridização de células somáticas. Fusão de diferentes tipos celulares.

Hibridização in situ. Método usado para determinar a localização cromossômica de um gene por outro fragmento específico de DNA, ou a distribuição de tecido de um mRNA ao usar uma sonda marcada que é complementar à sequência de interesse.

Hipermutação somática. Alta taxa de mutação somática como a que ocorre nos genes que codificam os anticorpos.

Hipótese de conflito genético. Sugere que o *imprinting* genômico evolui porque pressões diversas e conflitantes atuam nos alelos maternos e paternos para os genes que afetam o crescimento fetal. Por exemplo, os alelos paternos favorecem o crescimento fetal máximo, enquanto os alelos maternos favorecem menos esse crescimento máximo em função do elevado custo do crescimento fetal para a fêmea.

Hipótese de equilíbrio. Propõe que boa parte da variação molecular observada nas populações naturais é mantida ao equilibrar a seleção que favorece a variação genética.

Hipótese de Lyon. Proposta por Mary Lyon em 1961, essa hipótese propõe que um cromossomo X em cada célula feminina se torna inativado (um corpúsculo de Barr) e sugere que a determinação de qual cromossomo X se torna inativado é aleatória e varia entre células.

Hipótese de mutação neutra. Propõe que boa parte da variação molecular observada nas populações naturais é adaptativamente neutra e não afetada pela seleção natural. De acordo com esta hipótese, os indivíduos com variantes moleculares diferentes têm adaptabilidade igual.

Hipótese de um gene, um polipeptídio. Modificação da hipótese de um gene, uma enzima; propõe que cada gene codifica uma cadeia separada de polipeptídios.

Hipótese de um gene, uma enzima. Ideia proposta por Beadle e Tatum que cada gene codifica uma enzima separada.

Histona. Proteína de baixo peso molecular encontrada em eucariotos que se complexa com DNA para formar cromossomos.

Holoenzima. Complexo de uma enzima e outros fatores proteicos necessários para a função completa.

Homeobox. Subconjunto conservado dos nucleotídios nos genes homeóticos. Na *Drosophila*, é composto por 180 nucleotídios que codificam 60 aminoácidos de um domínio de ligação do DNA relacionado com o tema hélice-giro-hélice.

Homoplasmia. Presença de apenas uma versão de DNA no citoplasma de uma única célula.

Homozigose. Refere-se a um organismo que tem dois alelos idênticos em um *locus*.

Ilha CpG. Região do DNA com muitas cópias de uma base citosina seguida por uma base guanina; normalmente encontrada

próxima aos sítios de início de transcrição no DNA eucariótico. É comum as bases citosina nas ilhas CpG estarem metiladas quando os genes estão inativos, mas são desmetiladas antes do início da transcrição.

Imprinting genômico. Expressão diferenciada de um gene que depende do sexo do genitor que o transmitiu. A expressão do gene será diferente, dependendo se ele for herdado do genitor ou da genitora.

Imunidade celular. Tipo de imunidade gerada pelos linfócitos T, que reconhecem os antígenos encontrados na superfície das próprias células.

Imunidade humoral. Tipo de imunidade resultante de anticorpos produzidos pelos linfócitos B.

Indução coordenada. Síntese simultânea de várias enzimas que é estimulada por um único fator ambiental.

Indução. Estímulo da síntese de uma enzima por um fator ambiental, em geral a presença de um substrato específico.

Indutor. Substância que estimula a transcrição em um sistema induzível de regulação de gene; em geral, uma pequena molécula que se liga a uma proteína repressora e a altera de forma que o repressor não é mais capaz de se ligar ao DNA e inibir a transcrição.

Início abortado. Processo durante o início da transcrição no qual a RNA polimerase gera e libera repetidamente transcrições curtas, de 2 a 6 nucleotídios de comprimento, enquanto ainda está ligada ao promotor. Esse fenômeno ocorre em procariotos e eucariotos.

Inserção. Mutação na qual nucleotídios são adicionados a uma sequência de DNA.

Inserção no quadro de leitura. Inserção de alguns múltiplos de três nucleotídios, que não altera o quadro de leitura do gene.

Insulador. Sequência de DNA que bloqueia ou isola o efeito de um potencializador; deve estar localizada entre o potencializador e o promotor para ter sua atividade de bloqueio; também pode limitar o alcance das mudanças na estrutura da cromatina.

Integrase. Enzima que insere o DNA prófago, ou pró-viral, a um cromossomo.

Interação gênica. Interação entre genes em diferentes *loci* que afetam a mesma característica.

Interatoma. Conjunto completo de interações de proteína em uma célula.

Intercinese. Período entre a meiose I e a meiose II.

Interfase. Período no ciclo celular entre as divisões celulares. Na interfase, a célula cresce, se desenvolve e se prepara para a divisão celular.

Interferência. Grau no qual um *crossing over* interfere com *crossing overs* adicionais.

Interferência por RNA (RNAi). Processo no qual a clivagem do RNA de fita dupla produz pequenos RNA interferentes (siRNA) que se ligam ao mRNA contendo sequências complementares e levando à sua clivagem e degradação.

Intermediário de Holliday. Estrutura que se forma na recombinação homóloga; consiste em duas moléculas duplex conectadas por uma ponte cruzada.

Íntron. Sequência interveniente em um gene cortado; removido do RNA após a transcrição.

Íntron de RNA transportador (tRNA). Pertence a uma classe de íntrons nos genes do tRNA. A união desses genes depende de enzimas.

Íntron do grupo I. Pertence a uma classe de íntrons em alguns genes RNA ribossômicos capazes de fazer *auto-splicing*.

Íntron do grupo II. Pertence a uma classe de íntrons em alguns genes que codificam proteínas capazes de fazer *auto-splicing* e são encontrados em mitocôndrias, cloroplastos e algumas eubactérias.

Íntron pré-mRNA nuclear. Pertence a uma classe de íntron em genes que codificam proteínas que residem nos núcleos das células eucarióticas; removido por *splicing* mediado por spliceossomo.

Inversão cromossômica. Rearranjo no qual um segmento de um cromossomo foi invertido 180°.

Inversão paracêntrica. Inversão de cromossomo que não inclui o centrômero na região invertida.

Inversão pericêntrica. Inversão de cromossomo que inclui o centrômero na região invertida.

Inviabilidade híbrida. Mecanismo de isolamento reprodutivo no qual ocorre o acasalamento entre dois organismos de diferentes espécies e os descendentes híbridos produzidos não são viáveis.

Isolamento comportamental. Isolamento reprodutivo devido a diferenças no comportamento que evita o cruzamento.

Isolamento ecológico. Isolamento reprodutivo no qual espécies diferentes vivem em diversos *habitats* e interagem com o meio ambiente de variadas formas. Assim, seus membros não se encontram e não se reproduzem.

Isolamento gamético. Isolamento reprodutivo devido à incompatibilidade dos gametas. Ocorre acasalamento entre membros de diferentes espécies, mas os gametas não formam zigotos. É observado em muitas plantas, em que o pólen de uma espécie não pode fertilizar os óvulos de outra espécie.

Isolamento mecânico. Isolamento reprodutivo oriundo de diferenças anatômicas que evitam a cópula satisfatória.

Isolamento temporal. Isolamento reprodutivo no qual a reprodução de diferentes grupos ocorre em diversos momentos do ano, e então não existe fluxo de gene entre grupos; exemplificado por espécies de plantas que florescem em diferentes períodos do ano e, portanto, não trocam genes.

Isótopos. Diversas formas de um elemento com o mesmo número de prótons e elétrons, mas diferentes números de nêutrons no núcleo.

Junção de Holliday. Modelo de recombinação homóloga que é iniciada por rupturas da fita única em uma molécula de DNA.

Lei de Hardy-Weinberg. Princípio importante da genética de população que afirma que, em uma população de grande porte não afetada por mutação, migração ou seleção natural, que se acasala de forma aleatória, as frequências alélicas não mudam e as frequências genotípicas estabilizam após uma geração nas proporções p^2 (a frequência de *AA*), $2pq$ (a frequência de *Aa*) e q^2 (a frequência de *aa*), em que p é igual à frequência do alelo *A* e q é igual à frequência do alelo *a*.

Leptóteno. Primeiro subestágio da prófase I na meiose. No leptóteno, os cromossomos se contraem e se tornam visíveis.

Ligação completa. Ligação entre genes que estão próximos no mesmo cromossomo sem *crossing over* entre eles.

Ligação fosfodiéster. Ligação fosfodiéster unindo dois nucleotídios em uma cadeia de polinucleotídios.

Ligação incompleta. Ligação entre genes que exibem algum *crossing over*; intermediária nos seus efeitos entre a distribuição independente e a ligação completa.

Ligação peptídica. Ligação química que conecta os aminoácidos em uma proteína.

Ligador (linker). Pequeno segmento sintético de DNA que tem um ou mais sítios de restrição. Pode ser fixado pelas extremidades de qualquer pedaço de DNA e usado para inserir um gene em um vetor de plasmídio quando os locais de restrição não estão disponíveis.

LINE. Ver *Elemento longo disperso*.

Linfócito B. Tipo específico de linfócito que produz imunidade humoral; amadurece na medula óssea e produz anticorpos.

Linfócito T. Tipo específico de linfócito que produz imunidade humoral; origina-se na medula óssea e amadurece no timo.

Linhagem celular. Células idênticas, do ponto de vista genético, que se dividem indefinidamente e podem ser cultivadas em laboratório.

Locus. Posição em um cromossomo onde um gene específico está localizado.

Locus de traço quantitativo (QTL). Um gene ou região cromossômica que contribui para a expressão das características quantitativas.

Lod score (logaritmo de probabilidade). Logaritmo da razão da probabilidade de obter um conjunto de observações, considerando um grau específico de ligação, para a probabilidade de obter o mesmo conjunto de observações com a distribuição independente; usado para avaliar a probabilidade de ligação entre genes a partir de dados do heredograma.

Mapa físico. Mapa das distâncias físicas entre *loci*, marcadores genéticos ou outros segmentos de cromossomo; medido por pares de base.

Mapa genético (ligação). Mapa das distâncias relativas entre *loci* genéticos, marcadores ou outras regiões cromossômicas determinadas por taxas de recombinação; medidas em recombinação percentual ou unidades de mapa.

Mapeamento de deleção. Técnica para determinar a localização cromossômica de um gene ao estudar a associação de seu fenótipo ou produto com deleções cromossômicas particulares.

Mapeamento de restrição. Determina as localizações dos sítios cortados pelas enzimas de restrição em um pedaço de DNA.

Mapeamento intragênico. Mapeia as localizações de mutações dentro de um único *locus*.

Marcação. Método para adicionar um marcador radioativo ou químico em uma molécula como DNA ou RNA.

Marcador genético. Qualquer gene ou sequência de DNA usado para identificar uma localização em um mapa genético ou físico.

Marcadores epigenéticos. Mudanças herdáveis do DNA ou da estrutura da cromatina que não envolvem alteração na sequência

de bases e que produzem mudanças na expressão gênica.

Matriz nuclear. Rede de fibras de proteínas no núcleo; mantém o material nuclear no lugar.

Mecanismo de isolamento reprodutivo. Qualquer fator ou mecanismo biológico que evita a troca de genes.

Mecanismo de isolamento reprodutivo pós-zigótico. Isolamento reprodutivo que surge após a formação de um zigoto, porque os híbridos resultantes são inviáveis ou estéreis ou porque a reprodução é rompida nas gerações seguintes.

Mecanismo de isolamento reprodutivo pré-zigótico. Isolamento reprodutivo que evita a fusão dos gametas de duas espécies diferentes e, consequente, a formação de um zigoto híbrido.

Média. Medida estatística que descreve o centro de uma distribuição de medidas; calculada ao dividir a soma de todas as medidas pelo número de medidas; também chamada de média aritmética.

Mediador. Complexo de proteínas que é um dos componentes do aparelho basal de transcrição.

Megaesporo. Um dos quatro produtos da meiose nas plantas.

Megaesporócito. Uma célula reprodutora diploide presente no ovário das plantas, que sofre meiose para produzir macroesporos haploides.

Meio de cultura completo. Usado para cultivar bactérias ou outros microrganismos; contém todos os nutrientes necessários para crescimento e síntese, incluindo os que são normalmente sintetizados pelo organismo. As células com mutação para nutrientes podem crescer no meio completo.

Meio de cultura mínimo. Usado para cultivar bactérias ou algum tipo de microrganismo; tem apenas os nutrientes requeridos pelas células prototróficas (do tipo selvagem) – tipicamente, uma fonte de carbono, elementos essenciais como nitrogênio e fósforo, algumas vitaminas e outros íons e nutrientes necessários.

Meiose. Processo no qual os cromossomos de uma célula eucariótica se dividem para dar origem a células reprodutoras haploides. É composta por duas divisões: meiose I e meiose II.

Meiose I. Primeira fase da meiose. Na meiose I, o número de cromossomos é dividido pela metade.

Meiose II. Segunda fase da meiose. Os eventos que ocorrem na meiose II são semelhantes aos que ocorrem na mitose.

Merozigoto. Célula bacteriana que tem duas cópias de alguns genes, uma cópia do cromossomo bacteriano e uma segunda cópia de um plasmídio F introduzido, também chamado de diploide parcial. Metáfase. Um dos estágios da mitose. Na metáfase, os cromossomos se alinham no centro da célula.

Metáfase I. Um dos estágios da meiose I. Na metáfase I, os pares de cromossomos homólogos se alinham no centro da célula.

Metáfase II. Um dos estágios da meiose II. Na metáfase II, os cromossomos individuais se alinham na placa metafásica.

Metagenômica. Um campo emergente de tecnologia de sequenciamento no qual as sequências de genoma de um grupo de organismos que habitam um ambiente comum são amostradas e determinadas.

Metástase. Refere-se às células que se separam dos tumores malignos e se deslocam para outros locais, onde se estabelecem como tumores secundários.

Metilação do DNA. Modificação do DNA pela adição de grupos metila a algumas posições nas bases.

5-metilcitosina. Forma modificada da base citosina, contendo um grupo metila (CH_3) no carbono 5′.

5′-metilcitosina. Nucleotídio modificado, composto por citosina à qual foi adicionado um grupo metila, forma predominante de metilação no DNA eucariótico.

Microarranjo. Arranjo ordenado de fragmentos de DNA fixados em um suporte sólido, que serve como sondas para detectar sequências complementares; muitas vezes usado para avaliar a expressão de genes em vários tecidos e sob diferentes condições.

Microarranjo de proteínas. Grande número de proteínas diferentes aplicadas a uma lâmina de vidro como uma série de pontos, cada ponto com uma proteína diferente; usado para analisar interações entre proteínas.

Microbioma. Conjunto completo de todas as bactérias encontrada em um ambiente particular. Por exemplo, o microbioma humano é composto por todas as bactérias encontradas sob e dentro do corpo humano.

Microesporo. Produto haploide da meiose nas plantas.

Microesporócito. Célula reprodutora diploide no estame de uma planta; sofre meiose para produzir quatro microesporos haploides.

MicroRNA (miRNA). Pequena molécula de RNA, em geral com 21 a 22 pb de comprimento, produzida pela clivagem de RNA de fita dupla que surge a partir de pequenas estruturas em grampo (*hairpin*) dentro do RNA, que é principalmente de fita única.

Os miRNAs se combinam com proteínas para formar um complexo que se liga (imperfeitamente) a moléculas de mRNA e inibe sua tradução.

Microssatélite. Ver *Número variável de repetições em tandem* (*VNTRs*).

Microtúbulo. Fibra longa composta pela proteína tubulina; tem papel importante no movimento dos cromossomos na mitose e meiose.

Microtúbulo do fuso. Microtúbulo que move os cromossomos na mitose e meiose.

Migração. Movimento de genes de uma população para outra; também chamado de fluxo gênico.

Migração da ramificação. Movimento de uma ponte de cruzamento ao longo de duas moléculas de DNA.

Mitose. Processo pelo qual o núcleo de uma célula eucariótica se divide.

Modelo de ruptura de fita dupla. Modelo de recombinação homóloga no qual uma molécula de DNA sofre rupturas de fita dupla.

Modificação pós-tradução. Alteração de uma proteína após a tradução; pode incluir a clivagem a partir de uma proteína precursora maior, a remoção de aminoácidos e a fixação de outras moléculas na proteína.

Monofosfato de adenosina-3′,5′-cíclico (cAMP). Nucleotídio modificado que atua na repressão catabólica. Baixos níveis de glicose estimulam níveis elevados de cAMP; cAMP se prende a CAP, que se liga ao promotor de alguns óperons e estimula a transcrição.

Monossomia. Ausência de um cromossomo de um par de homólogos.

Morfógeno. Molécula cujo gradiente de concentração afeta o destino do desenvolvimento das células vizinhas.

Morgan. Cem (100) unidades de mapa.

Mosaicismo. Condição na qual as regiões de tecido dentro de um único indivíduo têm diferentes composições cromossômicas.

Motor molecular. Proteína especializada que move os componentes celulares.

mRNA policistrônico. Única molécula de RNA bacteriano que codifica mais de uma cadeia de polipeptídios; incomum nos eucariotos.

Mudança antigênica. Usada na referência a um vírus que sofreu modificações importantes no seu genoma por meio da redistribuição do material genético a partir de duas cepas diferentes do vírus.

Mutação. Mudança na informação genética hereditariamente transmitida.

Mutação adaptativa. Processo pelo qual um ambiente específico induz mutações que permitem que os organismos se adaptem ao ambiente.

Mutação ativadora (up). Mutação que aumenta a taxa de transcrição.

Mutação condicional. Mutação expressa apenas sob algumas condições.

Mutação constitutiva. Provoca a transcrição contínua de um ou mais genes estruturais.

Mutação cromossômica. Diferença a partir do tipo selvagem no número ou na estrutura de um ou mais cromossomos; em geral afeta muitos genes e tem grandes efeitos fenotípicos.

Mutação de ganho de função. Produz um novo traço ou faz um traço surgir em tecidos inadequados ou em períodos inadequados no desenvolvimento.

Mutação de perda de função. Leva à ausência completa ou parcial da função normal.

Mutação direta. Altera um fenótipo do tipo selvagem.

Mutação redutora. Reduz a taxa de transcrição.

Mutação condutora (driver). Uma mutação encontrada em uma célula de câncer que contribui para o processo do desenvolvimento da doença.

Mutação espontânea. Surge a partir de mudanças naturais na estrutura do DNA ou de erros na replicação.

Mutação no quadro de leitura (frameshift). Altera o quadro de leitura de um gene.

Mutação gênica. Afeta apenas um gene ou *locus*.

Mutação germinativa. Mutação em uma célula germinativa (que dá origem aos gametas).

Mutação induzida. Mutação consequente a agentes ambientais, como substâncias químicas ou radiação.

Mutação letal. Leva à morte prematura.

Mutação missense. Altera um códon no mRNA, resultando em um aminoácido diferente na proteína codificada.

Mutação neutra. Muda a sequência de aminoácidos de uma proteína, mas não altera a sua função.

Mutação nonsense. Mudança no códon específico de um aminoácido para um códon de parada.

Mutação passageira (passenger). Uma mutação encontrada em uma célula de câncer que não contribui para o processo do desenvolvimento da doença.

Mutação reversa (reversão). Mutação que muda um fenótipo mutante para o tipo selvagem.

Mutação silenciosa. Mudança na sequência de nucleotídios do DNA que não altera a sequência de aminoácidos de uma proteína.

Mutação somática. Mutação em uma célula que não dá origem a um gameta.

Mutação supressora. Mutação que esconde ou suprime o efeito de outra mutação em um sítio que é diferente do sítio da mutação original.

Mutação supressora intergênica. Mutação supressora que ocorre em um gene (*locus*) que é diferente do gene que tem a mutação original.

Mutação supressora intragênica. Mutação supressora que ocorre no mesmo gene (*locus*) que a mutação que o suprime.

Mutagênese direcionada por oligonucleotídio. Método de mutagênese direcionada pelo sítio que usa um oligonucleotídio para introduzir uma sequência mutante em uma molécula de DNA.

Mutagênese direcionada pelo sítio. Produz mudanças específicas de nucleotídio em sítios selecionados em uma molécula de DNA.

Mutágeno. Qualquer agente ambiental que aumente de forma significativa a taxa de mutação acima da taxa espontânea.

Não disjunção. Falha dos cromossomos homólogos ou cromátides-irmãs em se separar na meiose ou mitose.

Nó (ponto de ramificação). Ponto em uma árvore filogenética que representa um organismo. Os nós terminais são os ramos mais distantes da árvore e representam os organismos para os quais foram obtidos os dados. Os nós internos representam os ancestrais comuns aos organismos em diferentes ramos da árvore.

Northern blotting. Processo pelo qual o RNA é transferido de um gel para um suporte sólido como nitrocelulose ou filtro de náilon.

Notação. Ligação das informações da sequência de um gene que foi identificado com outras informações sobre a função e a expressão do gene, sobre a proteína codificada pelo gene e sobre genes semelhantes em outras espécies.

Nuclease modificada. Proteína consistindo de parte de uma enzima de restrição que cliva o DNA combinado com outra proteína que reconhece e se liga a uma sequência de DNA específica. Capaz de fazer cortes de fita dupla no DNA em uma sequência de DNA predeterminada. Essas nucleases podem ser projetadas para se ligar e cortar qualquer sequência específica de DNA.

Núcleo. Espaço nas células eucarióticas que é delimitado pelo envelope nuclear e tem os cromossomos.

Nucleoide. DNA bacteriano confinado em uma região definida do citoplasma.

Nucleosídio. Ribose ou desoxirribose ligada a uma base.

Nucleossomo. Unidade básica repetida da cromatina, composta por um cerne de oito histonas (duas de cada H2A, H2B, H3 e H4) e cerca de 146 pb de DNA que reveste o cerne duas vezes.

Nucleotídio. Unidade repetida de DNA ou RNA composta de um açúcar, um fosfato e uma base.

Nulissomia. Ausência de ambos os cromossomos de um par homólogo ($2n - 2$).

Número variável de repetições em tandem (VNTRs). Sequências curtas repetidas em tandem que variam muito em número entre os indivíduos, também chamada de microssatélites. Como são muito variáveis, são comumente usadas na identificação por DNA (DNA *fingerprinting*).

Oncogene. Gene dominante que estimula a divisão celular, levando à formação de tumores e contribuindo para o câncer, surge a partir de cópias mutantes de um gene celular normal (proto-oncogene).

Operador. Sequência de DNA no óperon de uma célula bacteriana. Uma proteína reguladora se liga ao operador e afeta a velocidade de transcrição dos genes estruturais.

Óperon. Conjunto de genes estruturais em uma célula junto com um promotor comum e outras sequências (como um operador)

que controla a transcrição dos genes estruturais.

Óperon induzível. Óperon ou outro sistema de regulagem gênica no qual a transcrição está desativada. Algo precisa acontecer para que a transcrição seja induzida ou iniciada.

Óperon reprimível. Óperon ou outro sistema de regulagem gênica no qual a transcrição está normalmente ativada. Algo precisa acontecer para que a transcrição seja reprimida ou desativada.

Organismo dioico. Pertence a uma espécie cujos membros têm ambas as estruturas reprodutivas masculina e feminina.

Organismo hermafrodito. Organismo que tem estruturas reprodutoras masculina e feminina.

Organismo-modelo da genética. Um organismo que é amplamente usado em estudos de genética em função de suas características, como tempo curto de geração e vários descendentes, que o tornam particularmente útil para análise genética.

Origem de replicação. Sequência de nucleotídios onde a replicação inicia.

Oscilação (wobble). Pareamento de base entre códon e anticódon no qual não há o pareamento padrão, em geral na terceira posição (3′) do códon; permite que mais de um códon pareie com o mesmo anticódon.

Overhang G. Uma cadeia de nucleotídios rica em guanina que se projeta além da cadeia complementar rica em C no final de um cromossomo.

Ovócito primário. Ovogônia que entrou na prófase I.

Ovócito secundário. Um dos produtos da meiose I nas fêmeas dos mamíferos; recebe a maior parte do citoplasma.

Ovogênese. Produção de óvulos em animais.

Ovogônia. Célula diploide no ovário; capaz de sofrer meiose para produzir um óvulo.

Óvulo. Gameta feminino; produto final da ovogênese.

Palíndromo. Sequência de nucleotídios que lê o mesmo nas fitas complementares; repetições invertidas.

Pangênese. Conceito inicial de hereditariedade que propõe que partículas carreiam informações genéticas oriundas de diferentes partes do corpo para os órgãos reprodutivos.

Papilomavírus humano (HPV). Vírus associado ao câncer de colo uterino.

Paquíteno. Terceiro subestágio da prófase I na meiose. O complexo sinaptonêmico se forma durante o paquíteno.

Par homólogo de cromossomos. Dois cromossomos que são semelhantes em estrutura e tamanho e carreiam informações genéticas para o mesmo conjunto de características hereditárias. Um cromossomo de um par homólogo é herdado da genitora e o outro é herdado do genitor.

Paradoxo do valor de C. A ausência de relação entre o tamanho do genoma e a complexidade nos organismos eucarióticos e a questão de qual é a função do DNA extra nas espécies com longos genomas.

Paramutação. Mudança epigenética na qual um alelo de um genótipo altera a expressão de outro alelo; a expressão alterada persiste por várias gerações, até após o gene alterado não estar mais presente.

PCR de transcrição reversa. Amplifica as sequências correspondentes ao RNA. A transcriptase reversa é usada para converter RNA em DNA complementar, que pode ser então amplificado pela reação em cadeia de polimerase.

PCR em tempo real. Modificação da reação em cadeia da polimerase que determina de forma quantitativa o ácido nucleico inicial; a quantidade de DNA amplificada é verificada à medida que a reação prossegue.

Penetrância incompleta. Refere-se a um genótipo que nem sempre expressa o fenótipo esperado. Alguns indivíduos têm o genótipo para um traço, mas não expressam o fenótipo.

Penetrância. Porcentagem de organismos com um determinado genótipo que expressa o fenótipo associado.

Pentaploidia. Organismo com cinco conjuntos haploides de cromossomos (5*n*).

Peptidil transferase. Atividade no ribossomo que cria uma ligação peptídica entre dois aminoácidos. A evidência sugere que essa atividade é feita por um dos componentes do RNA do ribossomo.

Perda da heterozigosidase. Em um *locus* com um alelo normal e outro mutante, inativação ou perda do alelo normal.

Pilus (pl. pili). Extensão da superfície de algumas bactérias que permite que ocorra conjugação. Quando um *pilus* em uma célula faz contato com um receptor em outra célula, o *pilus* se contrai e aproxima as duas células.

Pirimidina. Tipo de base nitrogenada no DNA e RNA. Citosina, timina e uracila são pirimidinas.

Placa. Mancha clara de células lisadas em uma camada contínua de bactérias na superfície do ágar de uma placa Petri. Cada

placa representa um único fago original que se multiplicou e lisou várias células.

Placa metafásica. Plano em uma célula entre dois polos de fuso. Na metáfase, os cromossomos individuais se alinham na placa metafásica.

Plasmídio. Pequena molécula de DNA circular encontrada nas células bacterianas que é capaz de se replicar independente do cromossomo bacteriano.

Plasmídio R (fator R). Plasmídio com genes que conferem resistência a antibióticos a qualquer célula que tenha esse plasmídio.

Plasmídio Ti. Grande plasmídio isolado da bactéria *Agrobacterium tumefaciens* e usado para transferir genes em células de plantas.

Pleiotropia. Um único genótipo influencia múltiplos fenótipos.

Pluripotência. Nas células-tronco embrionárias, a propriedade de ser indiferenciado, com a capacidade de formar todos os tipos de células em um organismo.

Polimorfismo de comprimento do fragmento de restrição (RFLP). Variação no padrão dos fragmentos produzidos quando moléculas de DNA são clivadas com a mesma enzima de restrição; representa uma diferença herdável nas sequências de DNA e pode ser usado nas análises genéticas.

Polimorfismo de único nucleotídio (SNP). Diferenças de único par de bases na sequência de DNA entre os membros de uma espécie.

Polimorfismo de único nucleotídio marcado (tag-SNP). Usado para identificar um haplótipo.

Polipeptídio. Cadeia de aminoácidos unidos por ligações peptídicas, também chamada de proteína.

Poliploidia. Ter mais de dois conjuntos haploides de cromossomos.

Polirribossomo. Molécula de RNA mensageiro com vários ribossomos presos.

Polo do fuso. Ponto a partir do qual os microtúbulos do fuso se irradiam.

Ponte cruzada. Em uma molécula de DNA heteroduplex, o ponto no qual cada cadeia de nucleotídios passa de uma molécula de DNA para outra.

Ponte dicêntrica. Estrutura produzida quando dois centrômeros de uma cromátide dicêntrica são “puxados” para polos opostos, alongando o cromossomo dicêntrico através do centro do núcleo. Consequentemente, a ponte dicêntrica se rompe à medida que os dois centrômeros são separados.

Ponto de ramificação. Nucleotídio de adenina nos íntrons do pré-mRNA nuclear que fica de 18 a 40 nucleotídios acima do sítio da união 3′.

Ponto de verificação G_1/S. Ponto importante no ciclo celular. Após passar pelo ponto de verificação em G_1/S, o DNA se replica e a célula está preparada para se dividir.

Ponto de verificação G_2/M. Ponto importante no ciclo celular próximo ao término de G_2. Após passar por esse ponto de verificação, a célula sofre mitose.

Ponto de verificação. Um ponto de transição importante no qual o progresso para o próximo estágio no ciclo celular é regulado.

Pool de genes. Total de genes em uma população.

População. O grupo de interesse; muitas vezes representado por um subconjunto chamado de amostra. Também um grupo de membros da mesma espécie.

População mendeliana. Grupo de indivíduos sexualmente aptos para reproduzir e cruzar.

Portador da translocação. Organismo heterozigoto para uma translocação cromossômica.

Pré-formacionismo. Conceito antigo de herança que propunha a existência de um adulto em miniatura (homúnculo) em cada óvulo ou espermatozoide e o aumento de seu tamanho no desenvolvimento, com todos os traços sendo herdados do genitor que fornece o homúnculo.

Primase. Enzima que sintetiza um trecho curto de RNA em um molde de DNA; atua na replicação para fornecer o grupo 3′-OH para a fixação de um nucleotídio de DNA.

Primeiro corpo polar. Um dos produtos da meiose I na ovogênese (oogênese); tem metade dos cromossomos, mas pouco do citoplasma.

Primer. Segmento curto de RNA ou DNA em um molde de DNA; fornece um grupo 3′-OH para a fixação de um nucleotídio de DNA no início da síntese de uma nova molécula de DNA na célula durante a replicação (*primer* de RNA) ou *in vitro* na reação da polimerase em cadeia ou PCR (*primer* de DNA).

Princípio da segregação independente (segunda lei de Mendel). Princípio importante da hereditariedade descoberto por Mendel que afirma que os genes que codificam características diferentes (genes em diferentes *loci*) se separam de forma independente; aplica-se apenas para genes localizados em diferentes cromossomos ou a genes distantes no mesmo cromossomo.

Princípio da segregação (primeira lei de Mendel). Princípio importante da hereditariedade descoberto por Mendel que afirma que cada indivíduo diploide tem dois alelos em um *locus* e que esses dois alelos se separam quando os gametas são formados, cada alelo indo para um gameta.

Princípio transformante. Substância responsável pela transformação. O DNA é o princípio transformante.

Probabilidade. Cálculo da ocorrência de um evento específico; mais formalmente, o número de vezes que um evento específico ocorre dividido pelo número de todos os possíveis desfechos. A probabilidade varia de 0 a 1.

Probabilidade condicional. Probabilidade que é modificada por informações adicionais de que outro evento ocorreu.

Probando. Uma pessoa com um traço ou doença para a qual foi construído um heredograma.

Procarioto. Organismo unicelular com estrutura celular única. Os procariotos incluem eubactérias e archaea.

Prófago. Genoma do fago que é integrado ao cromossomo bacteriano.

Prófase. Um dos estágios da mitose. Na prófase, os cromossomos se contraem e se tornam visíveis, o citoesqueleto se rompe e o fuso mitótico começa a se formar.

Prófase I. Estágio da meiose I. Na prófase I, os cromossomos se condensam e pareiam, ocorre o *crossing over*, a membrana nuclear se rompe e o fuso se forma.

Prófase II. Estágio da meiose após a intercinese. Na prófase II, os cromossomos se condensam, a membrana nuclear se rompe e o fuso se forma. Algumas células pulam esse estágio.

Projeto Proteoma Humano. Projeto com o objetivo de identificar e caracterizar todas as proteínas no corpo humano.

Pró-metáfase. Um dos estágios da mitose. Na pró-metáfase, a membrana nuclear é rompida e os microtúbulos do fuso se prendem aos cromossomos.

Promotor. Sequência de DNA à qual o aparelho de transcrição se liga para iniciar a transcrição; indica a direção da transcrição, qual das duas fitas de DNA será lida como molde e o ponto de início da transcrição.

Promotor interno. Promotor localizado dentro de sequências de DNA que são transcritas para o RNA.

Promotor regulador. Sequência de DNA localizada imediatamente acima do cerne do promotor que afeta a transcrição; tem sequências consenso às quais as proteínas reguladoras de transcrição se ligam.

Proteína alostérica. Proteína que muda sua conformação ao se ligar com outra molécula.

Proteína ativadora de transcrição. Proteína presente nas células eucarióticas que se liga a sequências consenso em promotores reguladores ou amplifica e afeta o início da transcrição ao estimular a montagem do aparelho basal de transcrição.

Proteína catabólica ativadora (CAP). Proteína que atua na repressão catabólica. Quando ligada ao cAMP, a CAP se liga ao promotor de alguns óperons e estimula a transcrição.

Proteína cromossômica não histona. Uma proteína oriunda da distribuição heterogênea das proteínas não histona na cromatina.

Proteína de choque térmico. Produzida por muitas células em resposta ao calor extremo e outros estresses; ajuda as células a evitar o dano causado por tais agentes de estresse.

Proteína de iniciação. Liga-se a uma origem da replicação e desenrola um pequeno segmento de DNA, tornando possível que a helicase e outras proteínas que se ligam à fita única se liguem e iniciem a replicação.

Proteína ligadora de DNA de fita única (SSB). Liga-se ao DNA de fita única na replicação e evita o anelamento com uma fita complementar e formação de estruturas secundárias.

Proteína ligadora de TATA (TBP). Cadeia de polipeptídios encontrada em diferentes fatores de transcrição que identifica e se liga a sequências em promotores eucarióticos.

Proteína de remodelagem da cromatina. A proteína se liga à sequência de DNA e rompe a estrutura da cromatina, fazendo com que o DNA se torne mais acessível à RNA polimerase e a outras proteínas.

Proteína ligadora de poli(A) (PABP). Liga-se à cauda poli (A) do mRNA eucariótico e o torna mais estável. Existem vários

tipos de PABP.

Proteína reguladora. Produzida por um gene regulador, uma proteína que se liga a outra sequência de DNA e controla a transcrição de um ou mais genes estruturais.

Proteínas SR. Grupo de proteínas ricas em serina e arginina que regulam a recomposição alternativa de pré-mRNA.

Proteinoquinase. Enzima que adiciona grupos de fosfato a outras proteínas.

Proteoma. Conjunto de todas as proteínas codificadas por um genoma.

Proteômica. Estudo do proteoma, o conjunto completo de proteínas encontrado em uma célula.

Proto-oncogene. Gene da célula normal que controla a divisão celular. Quando sofre mutação, pode se tornar um oncogene e contribuir para o progresso do câncer.

Pró-vírus. Cópia de DNA de DNA viral ou de RNA viral, integrado ao cromossomo do hospedeiro e replicado junto com esse cromossomo.

Pseudodominância. Expressão de um alelo normalmente recessivo graças a uma deleção no cromossomo homólogo.

Puff cromossômico. Intumescimento localizado de um cromossomo politênico, uma região da cromatina na qual o DNA está desenrolado e é submetido à transcrição.

Purina. Tipo de base nitrogenada no DNA e RNA. Adenina e guanina são purinas.

Quadrado de Punnett. Método de notação para determinar o desfecho de um cruzamento genético. Em uma grade, os gametas de um genitor são escritos ao longo da linha superior e os gametas do outro genitor são escritos ao longo da linha à esquerda. Dentro das células da grade, os alelos nos gametas são combinados para formar os genótipos dos descendentes.

Quadro de leitura. Forma de leitura de uma sequência de nucleotídios em grupos de três nucleotídios (códons) na tradução. Cada quadro de leitura começa com um códon de início e termina com um códon de parada.

Quadro de leitura aberta (ORF). Sequência contínua de nucleotídios de DNA que tem um códon de início e um códon de parada na mesma estrutura de leitura; assume-se que é um gene que codifica uma proteína, mas, em muitos casos, a proteína ainda não foi identificada.

Quiasma. Ponto de fixação entre os cromossomos homólogos no qual ocorre o *crossing over*.

Quinase dependente de ciclina (CDK). Uma proteína-chave no controle do ciclo celular; combina-se com a ciclina.

Ramo. Conexões evolutivas entre organismos em uma árvore filogenética.

Reação em cadeia da polimerase (PCR). Método para amplificar os fragmentos de DNA usando enzimas.

Reanelamento. Ver *Renaturação*.

Rearranjo cromossômico. Mudança do tipo selvagem na estrutura de um ou mais cromossomos.

Receptor de linfócito T. Encontrado na superfície de um linfócito T, um receptor que simultaneamente se liga a um antígeno estranho e ao próprio na superfície de uma célula.

Recessivo. Refere-se a um alelo ou fenótipo que é expresso apenas quando há duas cópias do alelo recessivo (homozigoto). O alelo recessivo não é expresso no fenótipo heterozigoto.

Recombinação. Processo que produz novas combinações de alelos.

Recombinação homóloga. Troca de informação genética entre moléculas de DNA homólogas.

Recombinação intracromossômica. Recombinação entre genes localizados no mesmo cromossomo.

Recombinação somática. Recombinação nas células somáticas, como os linfócitos maduros, entre os segmentos de genes que codificam anticorpos e receptores nos linfócitos T.

Recomposição alternativa. Ver *Splicing*.

Resgate genético. Introdução de nova variação genética em uma população endogâmica que melhora drasticamente a saúde dessa população em um esforço para aumentar suas chances de sobrevivência a longo prazo.

Região codificadora de proteína. A parte do mRNA composta por nucleotídios que especifica a sequência de aminoácidos de uma proteína.

Região codificadora de RNA. Sequência de nucleotídios de DNA que codifica uma molécula de RNA.

Região 3′ não traduzida (3′UTR). Sequências de nucleotídios na extremidade 3′ do mRNA; não codifica os aminoácidos de

uma proteína, mas afeta a estabilidade do mRNA e sua tradução.

Região 5′ não traduzida (5′UTR). Sequência de nucleotídios na extremidade 5′ do mRNA; não codifica os aminoácidos de uma proteína.

Região pseudoautossômica. Pequena região dos cromossomos X e Y com sequências de genes homólogos.

Regra de adição. Afirma que a probabilidade de que qualquer um de dois ou mais eventos mutuamente excludentes ocorra é calculada ao adicionar as probabilidades dos eventos individuais.

Regra de multiplicação. Afirma que a probabilidade de dois ou mais eventos independentes ocorrerem juntos é calculada ao multiplicar as probabilidades de cada um dos eventos individuais.

Regras de Chargaff. Regras desenvolvidas por Erwin Chargaff e seus colaboradores sobre as razões de bases no DNA.

Regressão. Análise de como uma variável muda em resposta a outra variável.

Regulação gênica. Mecanismos e processos que controlam a expressão fenotípica dos genes.

Relógio molecular. Refere-se ao uso de diferenças moleculares para estimar o tempo de divergência entre os organismos; pressupõe uma taxa quase constante na qual uma mutação neutra substitui outra.

Renaturação. Processo pelo qual duas moléculas de fita única de DNA complementares pareiam; também chamado de reanelamento.

Reparo de pareamento errado. Processo que corrige nucleotídios mal pareados no DNA após o término da replicação. As enzimas cortam os nucleotídios pareados errados da fita recém-sintetizada e usam a fita de nucleotídios original como molde quando fazem a substituição.

Reparo direto. Reparo do DNA no qual bases modificadas são trocadas por suas estruturas originais.

Reparo por excisão de bases. Reparo do DNA que primeiro retira as bases modificadas e então substitui o nucleotídio inteiro.

Reparo por excisão de nucleotídeo. Reparo de DNA que remove lesões volumosas e outros tipos de dano ao DNA.

Repetição curta em tandem (STR). Sequência de DNA muito curta repetida em *tandem* e amplamente encontrada no genoma humano.

Repetição direta flanqueadora. Sequência curta, repetida diretamente, produzida de cada lado de um elemento de transposição quando ele se insere no DNA.

Repetição expandida de nucleotídios. Mutação na qual o número de cópias de um conjunto de nucleotídios (a maioria de três nucleotídios) aumenta nas gerações seguintes.

Repetições invertidas. Sequências na mesma fita que estão invertidas e são complementares.

Repetições invertidas terminais. Sequências encontradas em ambas as extremidades de um elemento de transposição que são complementos invertidos de outra.

Replicação. Processo pelo qual o DNA é sintetizado a partir de um molde de nucleotídios de fita única.

Replicação bidirecional. Replicação de ambas as extremidades de uma bolha de replicação.

Replicação contínua. Replicação da fita principal no mesmo sentido que a fita desenrolada, tornando possível que novos nucleotídios sejam continuamente adicionados à extremidade 3′ da nova fita à medida que o molde é exposto.

Replicação de RNA. Processo em alguns vírus pelo qual o RNA é sintetizado a partir de um molde de RNA.

Replicação descontínua. Replicação da fita atrasada (*lagging*) no sentido oposto da fita que está desenrolada, o que significa que o DNA tem de ser sintetizado em pequenos intervalos (fragmentos de Okasaki).

Replicação por círculo rolante. Replicação de DNA circular que se inicia por uma ruptura em uma das fitas de nucleotídios, produzindo uma molécula de DNA circular de fita dupla e uma molécula de DNA linear de fita única, podendo a última circular servir como um molde para a síntese de uma fita complementar.

Replicação semiconservativa. Replicação na qual as duas fitas de nucleotídios do DNA se separam, cada uma servindo como molde para síntese de uma nova fita. Toda replicação do DNA é semiconservativa.

Replicação teta. Replicação de DNA circular que é iniciada pelo desenrolamento de duas fitas de nucleotídios, produzindo uma bolha de replicação. O desenrolar continua em uma ou ambas as extremidades da bolha, tornando-a progressivamente maior. A replicação do DNA em ambas as fitas do molde é simultânea com o desenrolar até que as duas forquilhas de replicação se encontrem.

Replicon. Unidade de replicação, composta por DNA a partir da origem da replicação para o ponto no qual a replicação de cada

lado da origem termina.

Repressão catabólica. Sistema de controle de genes em alguns óperons bacterianos no qual a glicose é usada preferencialmente e o metabolismo de outros açúcares é reprimido na presença de glicose.

Repressor. Proteína reguladora que se liga a uma sequência de DNA e inibe a transcrição.

Repulsão. Ver *Configuração trans.*

Resolvase. Enzima necessária para alguns tipos de transposição; responsável pela resolução, ou seja, o *crossing over* entre locais situados no elemento de transposição. A resolvase pode ser codificada pelo elemento de transposição ou por uma enzima celular que normalmente atua na recombinação homóloga.

Resposta à seleção. O quanto uma característica muda em uma geração devido à seleção; é igual à seleção diferencial vezes a herdabilidade de sentido restrito.

Resposta imunológica primária. Clone inicial de células específico para um antígeno específico e gerado quando o antígeno é encontrado pela primeira vez pelo sistema imunológico.

Resposta imunológica secundária. Clone de células gerado quando uma célula de memória encontra um antígeno, fornece imunidade duradoura.

Retrocruzamento. Cruzamento entre um indivíduo F_1 e um dos genótipos progenitores (P).

Retrotranspóson. Tipo de elemento de transposição nas células eucarióticas que tem algumas características dos retrovírus e transpõe através de um RNA intermediário.

Retrovírus. Vírus de RNA capaz de integrar seu material genético no genoma de seu hospedeiro. O vírus injeta seu genoma de RNA na célula hospedeira, onde a transcrição reversa produz uma molécula de DNA de fita dupla, complementar a partir do molde de RNA. A cópia de DNA então se integra ao cromossomo do hospedeiro para formar um pró-vírus.

Revisão. Capacidade das DNA polimerases em remover e substituir nucleotídios incorretamente pareados no curso da replicação.

Ribonucleoproteína nuclear pequena (snRNP). Estrutura encontrada nos núcleos das células eucarióticas composta por pequeno RNA nuclear (snRNA) e proteína; atua no processamento do pré-mRNA.

Ribonucleotídio. Nucleotídio contendo ribose; presente no RNA.

Ribose. Açúcar de cinco carbonos no RNA.

Riboswitch. Sequências reguladoras em uma molécula de RNA. Quando uma molécula indutora se liga ao *riboswitch*, a ligação muda a configuração da molécula de RNA e altera a expressão do RNA, em geral, ao afetar o término da transcrição ou afetar a tradução.

Ribozima. Molécula de RNA que pode agir como catalisador biológico.

RNA antisenso. Pequena molécula de RNA cuja base pareia com uma sequência de DNA complementar ou de RNA e afeta seu funcionamento.

RNA citoplasmático pequeno (scRNA). Pequena molécula de RNA encontrada no citoplasma das células eucarióticas.

RNA de interação com Piwi (piRNA). Pequena molécula de RNA que pertence a uma classe chamada de proteínas Piwi com as quais essas moléculas interagem; semelhante ao microRNA e ao pequeno RNA de interferência e acredita-se que participe na regulação do desenvolvimento do espermatozoide.

RNA guia (gRNA). Molécula de RNA que serve como molde para uma alteração feita no mRNA durante a edição do RNA.

RNA de interferência pequeno (siRNA). Molécula de RNA de fita única (em geral de 21 a 25 nucleotídios de comprimento) produzida por clivagem e processamento de RNA de fita dupla; liga-se a sequências complementares no mRNA e leva à clivagem e à degradação do mRNA. Alguns siRNAs ligam-se a sequências complementares no DNA e levam à sua metilação.

RNA mensageiro (mRNA). Molécula de RNA que carreia a informação genética para a sequência de aminoácidos de uma proteína.

RNA nuclear pequeno (snRNA). Pequena molécula de RNA encontrada nos núcleos das células eucarióticas; atua no processamento do pré-mRNA.

RNA nucleolar pequeno (snoRNA). Pequena molécula de RNA encontrada nos núcleos das células eucarióticas; atua no processamento do rRNA e na montagem dos ribossomos.

RNA polimerase. Enzima que sintetiza RNA a partir de um molde de DNA durante a transcrição.

RNA polimerase I. RNA polimerase eucariótica que transcreve grandes moléculas de RNA ribossômico (18 S rRNA e 28 S

rRNA).

RNA polimerase II. RNA polimerase eucariótica que transcreve RNA pré-mensageiro, alguns RNAs nucleares pequenos e alguns microRNAs.

RNA polimerase III. RNA polimerase eucariótica que transcreve RNA transportador, alguns RNAs ribossômicos pequenos (5 S rRNA), alguns RNAs nucleares e alguns microRNAs.

RNA polimerase IV. Transcreve pequenos RNAs interferentes nas plantas.

RNA polimerase V. Transcreve o RNA que participa na formação de heterocromatina nas plantas.

RNA pré-mensageiro (pré-mRNA). Molécula de RNA eucariótico que é modificada após a transcrição para se tornar mRNA.

RNA ribossômico (rRNA). Molécula de RNA que é o componente estrutural do ribossomo.

RNA transportador (tRNA). Molécula de RNA que carreia um aminoácido para o ribossomo e o transfere para a cadeia de polipeptídios crescente na tradução. É também chamado RNA de transferência.

RNA transportador-mensageiro (tmRNA). Uma molécula de RNA que tem propriedades de mRNA e tRNA; atua no resgate de ribossomos que estão parados na extremidade do mRNA.

Salto cromossômico. Método usado para mover de um gene em um fragmento clonado para sequências em fragmentos ligados a distância.

Segregação adjacente-1. Tipo de segregação que ocorre em um heterozigoto para uma translocação. Se os cromossomos originais não translocados são N_1 e N_2 e os cromossomos com os segmentos translocados são T_1 e T_2, então a segregação adjacente-1 ocorre quando N_1 e T_2 se movem para um polo e T_1 e N_2 se movem para o polo oposto.

Segregação adjacente-2. Tipo de segregação que ocorre em um heterozigoto para uma translocação. Se os cromossomos originais não translocados são N_1 e N_2 e os cromossomos com os segmentos translocados são T_1 e T_2, então a segregação adjacente-2 ocorre quando N_1 e T_1 se movem para um polo e T_2 e N_2 se movem para o polo oposto.

Segregação alternada. Tipo de segregação que ocorre em um heterozigoto para uma translocação. Se os cromossomos originais não translocados são N_1 e N_2 e os cromossomos com os segmentos translocados são T_1 e T_2, então a segregação alternada ocorre quando N_1 e N_2 se movem para um polo e T_1 e T_2 se movem para o polo oposto.

Segregação replicativa. Segregação aleatória de organelas nas células descendentes durante a divisão celular. Se houver duas ou mais versões de uma organela na célula original, o acaso determina a proporção de cada tipo que será separado em cada célula descendente.

Segundo corpo polar. Um dos produtos da meiose II na ovogênese (oogênese); tem um conjunto de cromossomos, mas pouco citoplasma.

Seleção artificial. Seleção praticada pelos humanos.

Seleção de célula fetal. Separação das células fetais a partir do sangue materno. Exame genético das células fetais que pode fornecer informações sobre doenças e distúrbios genéticos no feto.

Seleção diferencial. Diferença no fenótipo entre os indivíduos selecionados e a média da população inteira.

Seleção direcional. Seleção na qual um traço ou alelo é favorecido em relação a outro.

Seleção natural. Reprodução diferencial de genótipos.

Separase. Molécula que rompe as moléculas da coesina, que mantêm as cromátides-irmãs juntas.

Sequência associada a um telômero. Sequência encontrada nas extremidades de um cromossomo próxima à sequência telomérica; composta por longas sequências complexas repetidas.

Sequência centromérica. Uma sequência de DNA encontrada nos centrômeros funcionais.

Sequência consenso. Inclui os nucleotídios mais comumente encontrados em um local específico no DNA ou RNA.

Sequência consenso –10 (caixa de Pribnow). Sequência consenso (TATAAT) encontrada na maioria dos promotores bacterianos em aproximadamente 10 pb *upstream* do sítio de início da transcrição.

Sequência consenso –35. Sequência consenso (TTGACA) encontrada na maioria dos promotores bacterianos em aproximadamente 35 pb *upstream* do sítio de início da transcrição.

Sequência de inserção. Tipo único de elemento de transposição encontrado nas bactérias e seus plasmídios que tem apenas a informação necessária para seu próprio movimento.

Sequência de Shine-Dalgarno. Sequência consenso encontrada na região 5′ bacteriana não traduzida do mRNA; tem o sítio de ligação do ribossomo.

Sequência de replicação autônoma (ARS). Sequência de DNA que confere a capacidade de se replicar; tem uma origem de replicação.

Sequência repetida intercalada. Sequência repetida em múltiplas localizações por todo o genoma.

Sequência de sinalização. De 15 a 30 aminoácidos que são encontrados na extremidade amino de algumas proteínas eucarióticas e direcionam a proteína para localizações específicas na célula; em geral, rompida a partir da proteína.

Sequência telomérica. Sequência encontrada nas extremidades de um cromossomo; composta por várias cópias de sequências curtas e simples repetidas uma após a outra.

Sequenciamento baseado no mapa. Método de sequenciamento de um genoma no qual os fragmentos sequenciados são ordenados pela proximidade (*contigs*) com o uso de mapas genéticos ou físicos.

Sequenciamento de DNA. Processo que determina a sequência de bases ao longo de uma molécula de DNA.

Sequenciamento shotgun do genoma completo. Método de sequenciamento de genoma no qual os fragmentos sequenciados são montados na sequência correta em *contigs* ao usar apenas trechos de sequências sobrepostas (*overlaps*).

Sequências de repetição em tandem. Sequências de DNA repetidas uma após a outra tendem a se agrupar em localizações específicas em um cromossomo.

Sexo. Macho ou fêmea.

Sexo heterogamético. O sexo (masculino ou feminino) que produz dois tipos de gametas em relação aos cromossomos sexuais. Por exemplo, no sistema de determinação de sexo XX-XY, o macho produz gametas com X e com Y.

Sexo homogamético. O sexo (masculino ou feminino) que produz gametas que são semelhantes com relação aos cromossomos sexuais. Por exemplo, no sistema de determinação de sexo XX-XY, a fêmea produz gametas apenas com X.

Shelterina. Um complexo de multiproteínas que se liga a telômeros nos mamíferos e protege as extremidades do DNA de serem inadvertidamente reparadas como uma ruptura da fita dupla no DNA.

Silenciador. Sequência que tem muitas das propriedades presentes em um acentuador, mas reprime a transcrição.

Silenciamento do RNA. Mecanismo pelo qual o RNA de fita dupla é rompido e processado em pequenos RNA de interferência de fita única (siRNA), que se ligam a sequências complementares no mRNA e levam à clivagem e à degradação do mRNA; também conhecido como interferência por RNA e silenciamento pós-transcripcional por RNA do gene. Alguns siRNAs também se ligam a sequências complementares no DNA e guiam enzimas para metilar o DNA.

Sinapse. Pareamento próximo de cromossomos homólogos.

Síndrome de Down (trissomia do 21). Caracterizada por graus variáveis de déficit intelectual, traços faciais característicos, retardo no crescimento e no desenvolvimento e maior incidência de defeitos cardíacos, leucemia e outras anormalidades; causada pela duplicação de todo ou parte do cromossomo 21.

Síndrome de Down familiar. Causada por uma translocação robertsoniana na qual o braço longo do cromossomo 21 é translocado para outro cromossomo; tende a ocorrer em famílias.

Síndrome de Down primária. Causada pela existência de três cópias do cromossomo 21.

Síndrome de Edward (trissomia do 18). Síndrome caracterizada por grave déficit intelectual, orelhas de baixa implantação, pescoço curto, pés deformados, dedos das mãos cerrados, cardiopatia e outras incapacidades; é o resultado da presença de três cópias do cromossomo 18.

Síndrome de Klinefelter. Condição humana na qual as células têm um ou mais cromossomos Y junto com múltiplos cromossomos X (mais comumente XXY, mas também pode ser XXXY, XXXXY ou XXYY). As pessoas afetadas por essa síndrome têm aparência masculina, mas frequentemente com pequenos testículos, aumento das mamas e pelos, mais altos que o normal e estéreis, a maioria com inteligência normal.

Síndrome de Patau (trissomia do 13). Caracterizada por grave déficit intelectual, cabeça pequena, fronte olímpica, olhos pequenos, fenda labiopalatina, dedos extras nos pés e nas mãos e outras incapacidades; resulta da existência de três cópias do cromossomo 13.

Síndrome de Turner. Condição humana na qual as células têm apenas um cromossomo X e nenhum cromossomo Y (XO). As pessoas com a síndrome de Turner são do sexo feminino na aparência, mas não passam pela puberdade e apresentam características sexuais secundárias do sexo feminino pouco desenvolvidas; a maioria é estéril, mas com inteligência normal.

Síndrome do X frágil. Uma forma de déficit intelectual ligada ao X que aparece principalmente nos machos; surge a partir do aumento no número de repetições de um trinucleotídio CGG.

Síndrome do X triplo. Condição humana na qual as células têm três cromossomos X. Uma pessoa com essa síndrome tem um fenótipo feminino sem características diferentes além da tendência a ser alta e magra; algumas mulheres são estéreis, mas muitas menstruam normalmente e são férteis.

Sistema SOS. Sistema de proteínas e enzimas que permite que uma célula faça cópia do seu DNA quando há distorção na estrutura do DNA; produz vários erros na replicação e aumenta a taxa de mutação.

Sítio aminoacil (A). Um dos três sítios em um ribossomo ocupado por um tRNA na tradução. Todos tRNAs carregados (com a exceção do tRNA iniciador) entram primeiro no sítio A na tradução.

Sítio de início de transcrição. O primeiro nucleotídio do DNA que é transcrito em uma molécula de RNA.

Sítio de peptidil (P). Um dos três sítios em um ribossomo ocupado por um tRNA na tradução. No estágio de alongamento da síntese proteica, os tRNAs se movem do sítio do aminoacil (A) para o sítio P.

Sítio de saída (E). Um dos três locais em um ribossomo ocupado por um tRNA. No estágio de alongamento da tradução, o tRNA se move a partir do sítio do peptidil (P) para o sítio E, de onde ele sai do ribossomo.

Sítio de sequências marcadas (STS). Sequência de DNA curta (200 a 500 pb), presente apenas uma vez no genoma, cuja localização cromossômica foi determinada. Usada para determinar a localização genômica de um clone de DNA.

Sítio de recomposição do 3′. A extremidade 3′ de um íntron onde a clivagem ocorre na recomposição do RNA.

Sítio de recomposição do 5′. A extremidade 5′ de um íntron onde a segmentação ocorre na recomposição do RNA.

Sítio frágil. Constrição ou lacuna que aparece em um local específico em um cromossomo quando as células são cultivadas sob condições especiais. Um sítio frágil no cromossomo X humano está associado a déficit intelectual (síndrome do X frágil) e resulta de uma repetição expandida de trinucleotídios.

Sítio hipersensível da DNase I. Região da cromatina que se torna sensível à digestão pela enzima DNase I.

Sítios de clivagem 3′múltiplos. Refere-se à existência de mais de um sítio de clivagem 3′ em um único pré-mRNA, que torna possível que a clivagem e a poliadenilação ocorram em diferentes locais, produzindo mRNA de diferentes comprimentos.

Sobredominância. Seleção na qual o heterozigoto tem maior adaptabilidade do que o outro homozigoto; também chamada de vantagem do heterozigoto.

Sonda. Sequência conhecida de DNA ou RNA que é complementar a uma sequência de interesse e vai parear com esta; usada para encontrar sequências específicas de DNA.

Southern blotting. Processo pelo qual o DNA é transferido de um gel para um suporte sólido como nitrocelulose ou filtro de náilon.

Spliceossomo. Grande complexo composto por vários RNA e muitas proteínas que fazem *splicing* ou recomposição do pré-mRNA que codifica proteínas; tem cinco pequenas partículas de ribonucleoproteínas (U1, U2, U4, U5 e U6).

Splicing ou recomposição alternativa. Processo pelo qual um único pré-mRNA pode ser recomposto em mais de uma forma para produzir diferentes tipos de mRNA.

Splicing ou recomposição do RNA. Processo pelo qual íntrons são removidos e éxons são unidos.

Subdominância. Seleção na qual o heterozigoto tem menor adaptabilidade do que o outro homozigoto.

Substituição de base. Mutação na qual um único par de bases no DNA é alterado.

Subunidade ribossômica maior. A maior subunidade das duas subunidades de um ribossomo funcional.

Subunidade ribossômica menor. A menor subunidade das duas subunidades de um ribossomo funcional.

Superenrolamento. Estrutura terciária enrolada que se forma quando a fita é colocada em uma hélice de DNA ao enrolar excessivamente ou desenrolar a hélice. Um DNA superenrolado exibe superenrolamento positivo; um DNA desenrolado exibe superenrolamento negativo.

Superenrolamento negativo. Ver *Superenrolamento*.

Superenrolamento positivo. Ver *Superenrolamento*.

Tamanho efetivo da população. Número efetivo de adultos na idade fértil em uma população; influenciado pelo número de indivíduos que fornecem genes para a próxima geração, sua razão de sexos, variação entre indivíduos no sucesso reprodutivo, flutuações no tamanho da população, idade estrutural da população e se o acasalamento é aleatório.

Taq polimerase. DNA polimerase comumente usada nas reações de PCR. Isolada da bactéria *Thermus aquaticus*, a enzima é estável em altas temperaturas, e não é desnaturada durante a etapa de separação da fita do ciclo.

TATA box. Sequência consenso (TATAAAA) comumente encontrada em promotores eucarióticos da RNA polimerase II; em geral localizada a partir de 25 a 30 pb *upstream* do sítio de início da transcrição. A TATA *box* determina o ponto de início para a transcrição.

Taxa de mutação. Frequência de mudança de um gene do tipo selvagem para um mutante específico; em geral expresso como o número de mutações por unidade biológica (ou seja, mutações por divisão celular, por gameta ou por ciclo de replicação).

Tecnologia do DNA recombinante. Conjunto de técnicas moleculares para localizar, isolar, alterar, combinar e estudar segmentos de DNA.

Tecnologias de sequenciamento de próxima geração. Métodos de sequenciamento, como pirossequenciamento, que são capazes de determinar simultaneamente as sequências de vários fragmentos de DNA; as tecnologias são mais rápidas e mais baratas que o método de Sanger.

Telófase. Um dos estágios da mitose. Na telófase, os cromossomos chegam aos polos do fuso, o envelope nuclear é formado novamente e os cromossomos relaxam e se alongam.

Telófase I. Estágio da meiose I. Na telófase I, os cromossomos chegam aos polos do fuso.

Telófase II. Um dos estágios da meiose II. Na telófase II, os cromossomos chegam aos polos do fuso.

Telomerase. Enzima composta por proteína e RNA e que replica a extremidade (telômeros) dos cromossomos eucarióticos. A parte RNA da enzima tem um molde que é complementar a sequências repetidas no telômero e pareia com estas, fornecendo um molde para a síntese de cópias adicionais das repetições.

Telômero. Extremidade estável de um cromossomo.

Temperatura de desnaturação. Ponto central da faixa de desnaturação do DNA.

Teoria celular. Afirma que todas as formas de vida são compostas por células, que as células surgem apenas de outras células e que a célula é a unidade fundamental de estrutura e funcionamento nos organismos vivos.

Teoria cromossômica da herança. Afirma que os genes estão localizados nos cromossomos.

Teoria da seleção clonal. Explica a geração de respostas imunológicas primária e secundária. A ligação de um linfócito B a um antígeno estimula a célula a se dividir, dando origem a um clone de células geneticamente idênticas, todas específicas para o antígeno.

Teoria do germoplasma. Afirma que as células nos órgãos reprodutores carreiam um conjunto completo de informações genéticas.

Teoria endossimbiótica. Afirma que algumas organelas ligadas à membrana, como mitocôndrias e cloroplastos, nas células eucarióticas originaram-se como células eubacterianas de vida livre que entraram em uma relação endossimbiótica com uma célula hospedeira eucariótica e evoluíram para as organelas atuais; apoiadas por várias semelhanças de estrutura e sequência entre o DNA das organelas e DNA eubacteriano.

Terapia gênica. Uso de DNA recombinante para tratar uma doença ou distúrbio ao alterar a composição genética das células do paciente.

Terminador. Sequência de nucleotídios de DNA que leva ao término da transcrição.

Terminador dependente de Rho. Sequência no DNA bacteriano que exige a subunidade Rho da RNA polimerase para terminar a transcrição.

Terminador independente de Rho. Sequência no DNA bacteriano que não exige a subunidade Rho da RNA polimerase para terminar a transcrição.

Término da replicação. Ponto no qual a replicação para.

Teste de Ames. Teste no qual cepas especiais de bactérias são usadas para avaliar o potencial de substâncias químicas para causar câncer.

Teste de complementação. Teste projetado para determinar se duas mutações diferentes estão no mesmo *locus* (são alélicas) ou em diferentes *loci* (são não alélicas). Dois indivíduos que são homozigotos para duas mutações independentemente derivadas são cruzados, produzindo os descendentes F_1 que são heterozigotos para as mutações. Se as mutações estão no mesmo *locus*, F_1 terá um fenótipo mutante. Se as mutações estão em *loci* diferentes, F_1 terá um fenótipo do tipo selvagem.

Teste qui-quadrado de adequação de ajuste. Teste estatístico usado para avaliar como um conjunto de valores observados se

ajusta aos valores esperados. A probabilidade associada ao valor de qui-quadrado calculado é a probabilidade de que as diferenças entre os valores observados e esperados possam ser obra do acaso.

Teste de triagem no sangue materno. Teste para encontrar condições genéticas em um feto ao analisar o sangue da mãe. Por exemplo, o nível de α-fetoproteína no sangue materno fornece informações sobre a probabilidade de que um feto tenha um defeito do tubo neural.

Teste genético de coleta domiciliar. Teste para uma condição genética; o teste pode ser comparado diretamente pelo consumidor, sem o envolvimento de um médico ou outro profissional de saúde.

Tétrade. Os quatro produtos da meiose; todas as quatro cromátides de um par homólogo de cromossomos.

Tetraploidia. Possui quatro conjuntos haploides de cromossomos ($4n$).

Tetrassomia. Presença de duas cópias extras de um cromossomo ($2n + 2$).

Timina (T). Base pirimidina presente no DNA, mas não no RNA.

Tipo selvagem. O traço ou alelo que é mais comumente encontrado nas populações naturais (selvagens).

Topoisomerase. Enzima que adiciona ou remove rotações na hélice do DNA ao romper temporariamente as fitas de nucleotídios; controla o grau do superenrolamento do DNA.

Totipotência. O potencial de uma célula em se desenvolver em outro tipo de célula.

Tradução. Processo pelo qual uma proteína é montada a partir de informação contida no RNA mensageiro.

Transcrição. Processo pelo qual o RNA é sintetizado a partir de um molde de DNA.

Transcrição reversa. Síntese de DNA a partir de um molde de RNA.

Transcriptase reversa. Enzima capaz de sintetizar DNA complementar a partir de um molde de RNA.

Transcriptoma. Conjunto de todas as moléculas de RNA transcritas a partir de um genoma.

Transdução. Tipo de troca de genes que ocorre quando um vírus carreia genes de uma bactéria para outra. Após entrar na bactéria, o DNA recém-introduzido pode sofrer recombinação com o cromossomo bacteriano.

Transdução especializada. Transdução na qual os genes próximos de sítios especiais no cromossomo bacteriano são transferidos de uma bactéria para outra; requer bacteriófagos lisogênicos.

Transdução generalizada. Transdução na qual qualquer gene pode ser transferido a partir de uma célula bacteriana para outra por um vírus.

Transdutante. Célula bacteriana que recebeu genes de outra bactéria por meio da transdução.

Transesterificação. Reação química em algumas reações de recomposição (*splicing*) de RNA.

Transferência horizontal de genes. Transferência de genes de um organismo a outro por meio de mecanismo diferente da reprodução.

Transformação. Mecanismo pelo qual o DNA encontrado no meio é captado pela célula. Após a transformação, pode ocorrer recombinação entre os genes introduzidos e o cromossomo celular.

Transformante. Célula que recebeu material genético por meio da transformação.

Transgene. Gene estranho ou outro fragmento de DNA carreado no DNA da linhagem germinativa.

Transição. Substituição de base na qual uma purina é substituída por uma purina diferente ou uma pirimidina é substituída por uma pirimidina diferente.

Translocação. Movimento de um segmento do cromossomo para um cromossomo não homólogo ou para uma região no mesmo cromossomo. Também é o movimento de um ribossomo ao longo do mRNA durante a tradução.

Translocação não recíproca. Movimento de um segmento de cromossomo para um cromossomo ou região não homóloga sem qualquer (ou com desigual) troca recíproca de segmentos.

Translocação recíproca. Troca recíproca de segmentos entre dois cromossomos não homólogos.

Translocação robertsoniana. Translocação na qual os braços longos dos dois cromossomos acrocêntricos se unem em um centrômero comum, gerando um cromossomo com dois braços longos e normalmente outro cromossomo com dois braços curtos.

Transposase. Enzima codificada por vários tipos de elementos de transposição necessários para sua transposição. A enzima faz rupturas de fita única em cada extremidade do elemento de transposição e no lado da sequência-alvo onde o elemento se insere.

Transposição. Movimento de um elemento genético de transposição de um sítio para outro. A transposição replicativa aumenta

o número de cópias do elemento de transposição; a transposição não replicativa não aumenta o número de cópias.

Transposição não replicativa. Tipo de transposição na qual um elemento de transposição é cortado de um sítio antigo e se move para um novo local, resultando em nenhum aumento efetivo do número de cópias do elemento de transposição.

Transposição replicativa. Tipo de transposição na qual uma cópia do elemento de transposição se move para um novo lugar enquanto a cópia original permanece no sítio antigo; aumenta o número de cópias do elemento de transposição.

Transpóson composto. Tipo de elemento de transposição nas bactérias composto por duas sequências de inserção flanqueando um segmento de DNA.

Transpósons de DNA. Ver *Elemento de transposição.*

Trans-splicing. O processo de recomposição de éxons de dois ou mais pré-mRNAs.

Transversão. Substituição de base na qual uma purina é substituída por uma pirimidina ou vice-versa.

Triagem de heterozigoto. Testa os membros de uma população para identificar os heterozigotos portadores de um alelo responsável por uma doença que são saudáveis, mas com potencial para ter filhos com a doença.

Triagem de mutagênese. Método para identificar genes que influenciam um fenótipo específico. As mutações aleatórias são induzidas em uma população de organismos e são identificados os organismos com fenótipos mutantes. Esses organismos são cruzados para determinar a base genética do fenótipo e para mapear a localização das mutações responsáveis pelo fenótipo.

Triagem do recém-nascido. Testa recém-nascidos para alguns distúrbios genéticos.

Trifosfato de didesoxirribonucleosídio (ddNTP). Substrato especial para a síntese do DNA usado no método de sequenciamento didesoxi de Sanger, idêntico ao dNTP (o substrato comum para síntese de DNA), exceto por não ter um grupo 3′-OH. A incorporação de um ddNTP ao DNA termina a síntese deste.

Trifosfato de ribonucleosídio (rNTP). Substrato da síntese do RNA; composto por ribose, uma base nitrogenada, e três fosfatos ligados ao átomo de carbono 5′ da ribose. Na transcrição, dois dos fosfatos são clivados, produzindo um nucleotídio de RNA.

Triploidia. Ter três conjuntos haploides de cromossomos ($3n$).

Trissomia. Existência de uma cópia adicional de um cromossomo ($2n + 1$).

Trissomia do 8. Existência de três cópias do cromossomo 8; nos seres humanos, resulta em déficit intelectual, dedos dos pés e mãos contraídos, orelhas de baixa implantação malformadas e testa proeminente.

Trissomia do 13. Existência de três cópias do cromossomo 13; nos seres humanos, é responsável pela síndrome de Patau.

Trissomia do 18. Existência de três cópias do cromossomo 18; nos seres humanos, é responsável pela síndrome de Edward.

Trissomia do 21. Existência de três cópias do cromossomo 21; nos seres humanos, é responsável pela síndrome de Down.

tRNA carregado. Reação química na qual uma aminoacil-tRNA sintetase liga um aminoácido ao seu tRNA correspondente.

tRNA isoaceptores. Diferentes tRNA com diferentes anticódons que especificam o mesmo aminoácido.

Tubulina. Proteína encontrada nos microtúbulos.

Tumor maligno. Composto por células que são capazes de invadir outros tecidos.

Ultrassonografia. Procedimento para visualizar o feto. Um feixe de som de alta frequência é direcionado para o útero. Ondas sonoras que encontram tecido denso retornam e são transformadas em uma imagem do feto.

Unidade de mapa (m.u.). Unidade de medida para distâncias em um mapa genético; 1 unidade de mapa é igual a 1% de recombinação.

Unidade de transcrição. Sequência de nucleotídios no DNA que codifica uma única molécula de RNA, junto com as sequências necessárias para sua transcrição; normalmente tem um promotor, uma sequência codificadora de RNA e um terminador.

Uracila (U). Base pirimidina existente no RNA, mas não normalmente no DNA.

Valor de C. Quantidade haploide de DNA encontrada em uma célula de um organismo.

Variação antigênica. Usada na referência a vírus que evoluem rapidamente, com novas cepas surgindo com frequência por causa das mutações.

Variação no número de cópias (CNV). Diferença entre organismos no número de cópias de qualquer sequência grande de DNA (maior que 1.000 pb).

Variância. Medida estatística que descreve a variabilidade de um grupo de medidas.

Variância ambiental. Componente da variância fenotípica devido a diferenças ambientais entre membros de uma população.

Variância da interação genética ambiental. Componente da variância fenotípica que surge a partir de uma interação entre genótipo e ambiente. Os genótipos são expressos de forma diferente em diferentes ambientes.

Variância da interação gênica. Componente da variância genética que pode ser atribuído à interação gênica (interação entre genes em diferentes *loci*).

Variância fenotípica. Mede o grau de diferenças fenotípicas entre um grupo de indivíduos; composta por variâncias genéticas, ambientais e de interação genético-ambiental.

Variância genética. Componente da variância fenotípica devido a diferenças ambientais entre membros de uma população.

Variância genética aditiva. Componente da variância genética que pode ser atribuído ao efeito aditivo de diferentes genótipos.

Variância genética por dominância. Componente da variância genética que pode ser atribuído à dominância (interação entre genes no mesmo *locus*).

Vetor de clonagem. Molécula de DNA estável, replicativa, à qual um fragmento de DNA estranho pode ser preso e transferido para uma célula hospedeira.

Vetor de expressão. Vetor de clonagem contendo sequências de DNA como um promotor, um sítio de ligação de ribossomo e sítios de iniciação e terminação de transcrição, que permitem que fragmentos de DNA inseridos no vetor sejam transcritos e traduzidos.

Vetor de transferência. Vetor de clonagem que permite que o DNA seja transferido para mais de um tipo de célula hospedeira.

Via de processamento alternativa. Uma das diversas vias pelas quais um único pré-mRNA pode ser processado de várias formas para produzir tipos alternativos de mRNA.

Via de transdução de sinal. Sistema no qual um sinal externo (iniciado por um hormônio ou fator de crescimento) deflagra uma cascata de reações intracelulares que, por fim, produz uma resposta específica.

Vigilância do mRNA. Mecanismos para detecção e eliminação de mRNAs com erros que podem criar problemas no curso da tradução.

Vírus. Agente replicativo não celular composto por ácido nucleico circundado por um revestimento proteico; só pode se replicar dentro da célula hospedeira.

Western blotting. Processo pelo qual a proteína é transferida de um gel para um suporte sólido como nitrocelulose ou filtro de náilon.

Z-DNA. Estrutura secundária do DNA caracterizada por 12 bases por giro, uma hélice girando para esquerda e um esqueleto de açúcar-fosfato que faz zigue-zague para frente e para trás.

Zigóteno. Segundo subestágio da prófase I na meiose. No zigóteno, os cromossomos entram na sinapse.

Respostas a Questões e Problemas Selecionados

Capítulo 1

1. Na cultura Hopi, as pessoas com albinismo eram consideradas especiais e recebiam *status* especial. Como a intensa exposição à luz do sol poderia ser prejudicial ou mortal, os homens albinos não trabalhavam na agricultura. O albinismo era considerado um traço positivo, em vez de uma condição física negativa, permitindo que as pessoas albinas tivessem mais crianças e aumentassem a frequência do alelo. Finalmente, a pequena população da tribo Hopi pode ter ajudado a aumentar a frequência aleatória do alelo do gene albino.

17. A evolução é uma mudança genética ao longo do tempo. Para que ocorra a evolução, primeiro é necessário que haja a variação genética e então as forças evolutivas modificam a proporção das variantes genéticas ao longo do tempo. Assim, a variação genética é a base de toda mudança evolutiva.

18. **(a)** Genética da transmissão; **(b)** genética de populações; **(c)** genética de poluações; **(d)** genética molecular; **(e)** genética molecular; **(f)** genética da transmissão.

22. **(a)** Teoria do germoplasma; **(b)** Pré-formacionismo; **(c)** Herança das características adquiridas **(d)** Pangênese.

23. **(a)** A pangênese postula que partículas específicas (chamadas de gêmulas) carreiam informações genéticas de todas as partes do corpo para os órgãos reprodutores e que essas informações são transmitidas para o embrião. No embrião cada unidade direciona a formação de sua própria parte específica do corpo. De acordo com a teoria do germoplasma, as células produtoras de gameta encontradas nos órgãos reprodutores têm um conjunto de informações genéticas que é transmitido para os gametas. Os dois conceitos são semelhantes, pois propõem que as informações genéticas estão em unidades distintas, transferidas para os descendentes. A diferença entre os dois conceitos consiste na localização das informações genéticas. Na pangênese, estão localizadas em diferentes partes do corpo e têm de ser levadas para os órgãos reprodutores. Na teoria do germoplasma, toda informação genética já está nas células reprodutoras.

    (b) O pré-formacionismo afirma que o espermatozoide ou o óvulo tem uma miniatura pré-formada adulta chamada de homúnculo. No desenvolvimento, o homúnculo cresce para produzir um descendente. Apenas um genitor contribui com os traços genéticos para o descendente. A herança combinada demanda contribuições de material genético de ambos os genitores. As contribuições genéticas dos genitores se misturam para produzir o material genético dos descendentes. Uma vez misturado, o material genético não pode ser separado nas futuras gerações.

    (c) A herança das características adquiridas postula que os traços adquiridos na vida de uma pessoa alteram o material genético e podem ser transmitidos para os descendentes. Nossa moderna teoria de hereditariedade indica que os descendentes herdam genes localizados nos cromossomos transmitidos por seus genitores. Estes cromossomos se separam na meiose nas células germinativas dos genitores e são transmitidos para os gametas.

24. **(a)** Ambos os tipos celulares têm membranas lipídicas bipolares, genomas de DNA e maquinário para replicação, transcrição, tradução do DNA, metabolismo energético, resposta aos estímulos, crescimento e reprodução. As células eucarióticas têm um núcleo com DNA cromossômico e organelas internas ligadas à membrana.

    (b) Um gene é uma unidade básica de informação hereditária, em geral codifica um RNA funcional ou um polipeptídio. Os alelos são formas variantes de um gene e surgem graças à mutação.

    (c) O genótipo é o conjunto de genes ou alelos herdados por um organismo de seu(s) genitor(es). A expressão dos genes de um genótipo específico, por interação com fatores ambientais, produz o fenótipo, o traço observável.

    (d) Ambos são polímeros de ácidos nucleicos. O RNA tem um açúcar ribose, enquanto o DNA tem um açúcar desoxirribose. O RNA também tem uracila como uma de suas quatro bases, enquanto o DNA tem timina. As outras três bases são comuns ao DNA e RNA. Finalmente, o DNA é bifilamentar, composto por duas fitas complementares, enquanto o RNA é unifilamentar (fita única).

    (e) Os cromossomos são estruturas compostas por DNA e proteínas associadas. O DNA contém as informações genéticas.

25.

Tipo de albinismo	**Fenótipo**	**Gene mutante**
OCA2	Pigmento reduzido na pele, no cabelo e nos olhos, mas pouco pigmento adquirido com idade;	OCA2

	problemas visuais	
OCA1B	Ausência geral de pigmento no cabelo, na pele, nos olhos, mas pode ter um pouco de pigmento, não varia com a idade; problemas visuais	Tirosinase
OCA1A	Completa ausência de pigmento; problemas visuais	Tirosinase
OCA3	Existe algum pigmento, mas sensibilidade ao sol e problemas visuais	Proteína 1 relacionada à tirosinase
OASD	Ausência de pigmento nos olhos e surdez posteriormente	Desconhecido
OA1	Ausência de pigmento nos olhos, mas pigmento normal em outras partes	Gene GPR143
ROCA	Coloração cobre-claro na pele e no cabelo de africanos; diluição da cor da íris	Proteína 1 relacionada à tirosinase
OCA4	Pigmentação reduzida	MATP

27. Todos os genomas têm necessariamente a capacidade de armazenar informações complexas e variar. O diagrama para um organismo inteiro precisa estar contido no genoma de cada célula reprodutiva. As informações têm de estar codificadas de modo que possam ser usadas como um conjunto de instruções para montar os componentes das células. O material genético de qualquer organismo precisa ser estável, replicado com precisão e transmitido fidedignamente para os descendentes e tem de ser capaz de sofrer mutação.

Capítulo 2

20. **(a)** As duas cromátides de um cromossomo.

    (b) Os dois cromossomos de um par homólogo.

    (c) Coesina.

    (d) A enzima separase.

    (e) As mãos dos dois homens cegos.

    (f) Se um homem não consegue pegar sua meia, seria difícil usar a faca para cortar o fio que as mantêm unidas. As duas meias de um par não estariam separadas e ambas ficariam na bolsa de um dos homens. Da mesma forma, se cada cromátide não estiver presa às fibras do fuso e puxada em sentidos opostos, as duas cromátides não se separarão e ambas migrarão para a mesma célula. Esta célula tem duas cópias de um cromossomo.

23.

Estágio	Número de células contadas	Proporção de células em cada estágio	Duração média (horas)
Interfase	160	0,80	19,2
Prófase	20	0,10	2,4
Pró-metáfase	6	0,03	0,72
Metáfase	2	0,01	0,24
Anáfase	7	0,035	0,84
Telófase	5	0,025	0,6
Total	200	1,0	24

A duração média da fase M pode ser determinada ao adicionar as horas gastas em cada estágio da mitose. Nestas células, a fase M dura 4,8 horas. A metáfase requer 0,24 hora ou 14,4 minutos.

26. **(a)** 12 cromossomos e 24 moléculas de DNA; **(b)** 12 cromossomos e 24 moléculas de DNA; **(c)** 12 cromossomos e 24 moléculas de DNA; **(d)** 12 cromossomos e 24 moléculas de DNA; **(e)** 12 cromossomos e 12 moléculas de DNA; **(f)** 6 cromossomos e 12 moléculas de DNA; **(g)** 12 cromossomos e 12 moléculas de DNA; **(h)** 6 cromossomos e 6 moléculas de DNA.

28. **(a)** O número diploide de cromossomos é seis; **(b)** a célula do lado esquerdo está na anáfase I da meiose; a célula do meio está na anáfase da mitose; a célula do lado direito está na anáfase II da meiose; **(c)** a célula do lado esquerdo possui 6 cromossomos e 12 moléculas de DNA; a célula do meio possui 12 cromossomos e 12 moléculas de DNA; a célula do lado direito possui 6 cromossomos e 6 moléculas de DNA.

29.

Estágio da mitose		Quantidade de DNA por célula
a. G_1	d	3,7 pg
b. Prófase I	a, f	7,3 g
c. G_2	b, c, e	14,6 pg
d. Após telófase II e citocinese		
e. Anáfase I		
f. Metáfase II		

A quantidade de DNA na célula será dobrada após o término da fase S no ciclo celular e antes da citocinese tanto na mitose quanto na meiose I. Ao término da citocinese após a meiose II, a quantidade de DNA estará dividida à metade.

a. G_1 ocorre antes da fase S e da duplicação da quantidade de DNA.

b. A quantidade de DNA duplica na fase S; portanto, durante a prófase I da meiose, a quantidade de DNA na célula é o dobro que a quantidade em G.

c. G_2 ocorre logo após o término da fase S, então a quantidade de DNA é o dobro que a quantidade na G_1.

d. Após a citocinese associada com a meiose II, cada célula-filha terá apenas metade da quantidade de DNA de uma célula encontrada em G_1 da interfase.

e. Durante a anáfase I da meiose, a quantidade de DNA na célula é o dobro que em G_1.

f. Na metáfase II da meiose, a quantidade de DNA em cada célula é a mesma que na fase G_1 porque o DNA dobrou na fase S, mas foi então reduzida à metade na primeira divisão meiótica.

30. **(a)** Se a coesina não se formar no início da mitose, as cromátides-irmãs podem separar-se antes da anáfase. O resultado seria a separação inadequada dos cromossomos para as células-filhas.

(b) Se não houver shugosina durante a meiose, então a coesina no centrômero pode ser degradada, permitindo a separação das cromátides-irmãs junto com os homólogos durante a anáfase I, levando a separação inadequada dos cromossomos para as células-filhas.

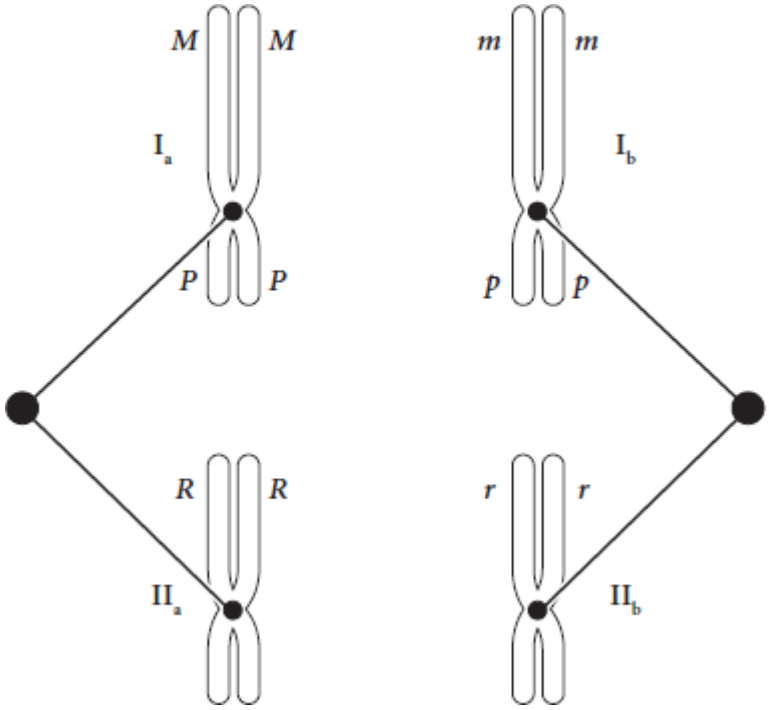
(c) Se a shugoshina não for degradada, a coesina no centrômero permanecerá protegida da degradação. A coesina intacta evitará que as cromátides-irmãs se separem durante a anáfase II da meiose, levando a uma separação inadequada das cromátides-irmãs e células-filhas com muito ou pouco cromossomo.

(d) Os cromossomos homólogos e as cromátides-irmãs não se separam na meiose e mitose, resultando em algumas células com muito pouco cromossomo e algumas células com cromossomo em excesso.

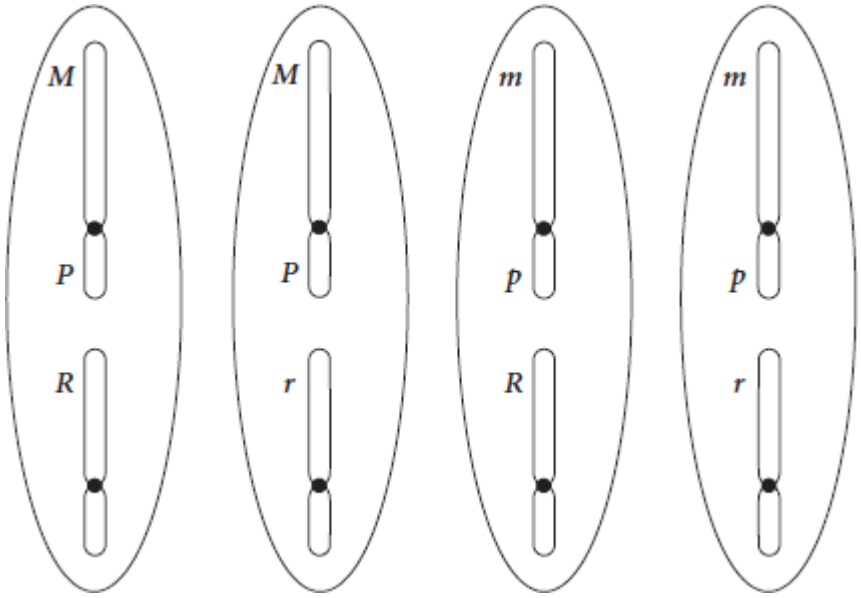
31. Uma célula do mesmo organismo que está na prófase da mitose teria 24 cromossomos. Durante a prófase I da meiose a célula terá 24 cromossomos.

33. Espera-se que os descendentes de um organismo com o número maior de pares homólogos de cromossomos exibam maior variação. O número de diferentes combinações de cromossomos que são possíveis no gameta é $2n$, em que n é igual ao número de pares homólogos de cromossomos. Para a mosca-da-fruta, que tem quatro pares de cromossomos, o número de possíveis combinações é $2^4 = 16$. Para a mosca-doméstica, que tem seis pares de cromossomos, o número de possíveis combinações é $2^6 = 64$.

34. **(a)** Metáfase I



(b) Gametas



36.

	Tipo de célula	Número de cromossomos	Número de moléculas de DNA
a.	Espermatogônia	64	64

Assumindo que a espermatogônia está na fase G_1 antes da produção das cromátides-irmãs na fase S, o número de cromossomos será o número diploide de cromossomo.

b.	Primeiro corpo polar	32	64

O primeiro corpo polar é o produto da meiose I então ele será haploide, mas as cromátides-irmãs não estão separadas, então cada cromossomo terá duas cromátides-irmãs.

c.	Ovócito primário	64	128

O ovócito primário parou na prófase I da meiose. Então os homólogos não se separaram ainda e cada cromossomo tem duas cromátides-irmãs.

d.	Espermatócito secundário	32	64

O espermatócito secundário é o produto da meiose I e ainda vai entrar na meiose II. Então o espermatócito secundário

será haploide porque os pares homólogos foram separados na meiose I, mas cada cromossomo ainda é composto de duas cromátides-irmãs.

38. **(a)** Não, a informação não é idêntica à encontrada no ovócito secundário. O primeiro corpo polar e o ovócito secundário são o resultado da meiose I, que produz duas células não idênticas. O primeiro corpo polar e o ovócito secundário terão apenas um membro de cada par de cromossomos original. Além disso, o *crossing over* que ocorre na prófase I vai gerar novos e diferentes arranjos do material genético nestas cromátides que participaram do *crossing over*.

    (b) Não, a informação não é idêntica. O segundo corpo polar e o óvulo terão cópias dos mesmos membros dos pares homólogos de cromossomos que são separados na meiose. Entretanto, por causa do *crossing over*, as cromátides-irmãs que se separam na anáfase II e dão origem ao óvulo e ao segundo corpo polar não são mais idênticas.

41. A maioria dos machos produz espermatozoide por meiose. Como a meiose ocorre apenas nas células diploides, os machos haploides das abelhas não sofrem meiose. As abelhas-machos podem produzir espermatozoides, mas apenas por mitose. As células haploides que se dividem por mitose produzem células haploides.

Capítulo 3

13. Características úteis: são fáceis de criar e manter, crescem rapidamente e produzem várias gerações em pouco tempo. Exemplos de organismos: *Neurospora*, um fungo; *Saccharomyces cerevisiae*, uma levedura; *Arabidopsis*, uma planta; *Caenorhabditis elegans*, um nematódeo; *Drosophila melanogaster*, uma mosca-da-fruta.

17. **(a)** Embora a fêmea branca tenha descendentes, seus ovos foram produzidos pelo ovário da porquinha-da-índia fêmea preta. O ovário transplantado produziu apenas os óvulos com o alelo para a cor preta do pelo. Como a maioria dos mamíferos, as fêmeas do porquinho-da-índia produzem ovócitos primários no início do desenvolvimento e assim o ovário transplantado já tinha ovócitos primários produzidos pela fêmea de cor preta.

    (b) O macho de cor branca contribuiu com o alelo *w*, enquanto a fêmea de cor branca contribuiu com o alelo *W* do ovário transplantado. Desta forma, os descendentes são *Ww*.

    (c) A produção de descendentes de porquinho-da-índia de cor preta sugere que o alelo para a cor preta do pelo foi transmitido para os descendentes a partir do ovário transplantado de acordo com a teoria do germoplasma. Se a pangênese estivesse correta, os descendentes do porquinho-da-índia seriam brancos. Os alelos para pelo branco teriam se deslocado para o ovário transplantado e então para os gametas da fêmea branca. A ausência de descendentes brancos indica que não ocorreu pangênese.

18. **(a)** A genitora é i^Bi^B, o genitor é I^AI^B.

    (b) Ambos os genitores são i^Bi^B.

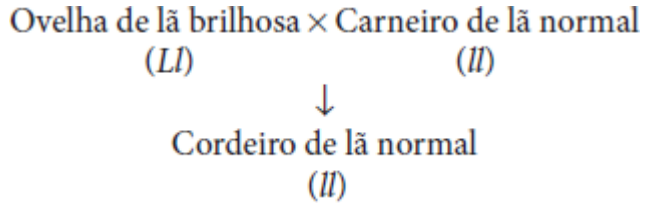
    (c) O genitor é i^Bi^B, a genitora é I^AI^A ou possivelmente I^Ai^B, mas é improvável que uma fêmea heterozigota neste cruzamento tenha produzido oito filhotes do tipo sanguíneo A por acaso apenas.

    (d) Ambos os genitores são I^Ai^B.

    (e) Ambos os genitores são I^AI^A, ou um genitor é I^AI^A e o outro genitor é I^Ai^B. O tipo sanguíneo dos descendentes não permite uma determinação precisa do genótipo de cada genitor.

    (f) A genitora é i^Bi^B, o genitor é I^Ai^B.

21. Sim, é possível determinar o genótipo de cada genitor, considerando que o alelo brilhoso dominante (*L*) exibe dominância completa. O carneiro e o único cordeiro têm de ser homozigotos para o alelo normal (*l*) porque ambos têm o genótipo de lã normal. Como o cordeiro recebe apenas um único alelo (*l*) do carneiro, a ovelha necessariamente contribuiu com o outro alelo *l* recessivo. Portanto, a ovelha tem de ser heterozigota para o alelo brilhoso. Em resumo:



22. **(a)** Sally (*Aa*), mãe de Sally (*Aa*), pai de Sally (*aa*), e irmão de Sally (*aa*); **(b)** 1/2; **(c)** 1/2.

25. **(a)** Porque o 2 é encontrado em um lado de um dado de 6 lados; então, existe uma chance de 1/6 de jogar um dois.

    (b) A probabilidade de conseguir 1 em um dado de 6 lados é 1/6. Da mesma forma, a probabilidade de conseguir 2 em

um dado de 6 lados é 1/6. Como a pergunta é *sobre a probabilidade de conseguir 1 ou 2*, e estes eventos são mutuamente exclusivos, devemos usar a regra aditiva de probabilidade para determinar a probabilidade de jogar 1 ou 2: (*p* de conseguir 1) + (*p* de conseguir 2) = *p* de conseguir 1 ou 2 = 1/6 + 1/6 = 2/6 = 1/3 de jogar um 1 ou um 2.

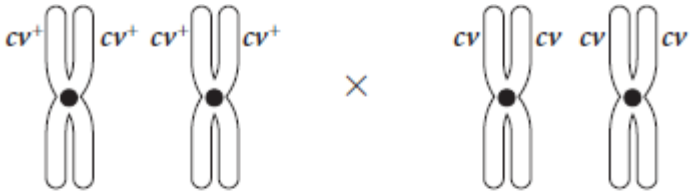
(c) A probabilidade de conseguir acertar um número par depende do número de números pares encontrados no dado. Um dado simples tem 3 números pares (2, 4, 6). A probabilidade de conseguir um destes 3 números em um dado de 6 lados é 1/6. Para determinar a probabilidade de conseguir um 2, um 4 ou um 6, aplicamos a regra aditiva: 1/6 + 1/6 + 1/6 = 3/6 = 1/2.

(d) O número 6 é encontrado apenas em um lado do dado de 6 lados. Portanto, a probabilidade de acertar um 6 é 1/6. A probabilidade de jogar qualquer número, exceto 6 é (1 – 6/6) = 5/6.

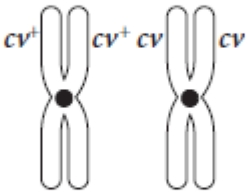
26. **(a)** 1/18; **(b)** 1/36; **(c)** 11/36; **(d)** 1/6; **(e)** 1/4; **(f)** 3/4.

27. **(a)** 1/128; **(b)** 1/64; **(c)** 7/128; **(d)** 35/128; **(e)** 35/128.

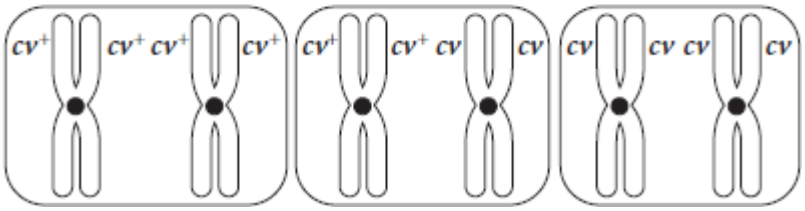
29. Genitores:



Geração F_1:



Geração F_2:



30. **(a)** Nos porquinhos-da-índia pretos F_1 (*Bb*), apenas um cromossomo tem o alelo preto, então o número de cópias presentes em cada estágio é: G_1, um alelo preto; G_2, dois alelos pretos, metáfase da mitose, dois alelos pretos, metáfase I da meiose, dois alelos pretos, após a citocinese da meiose, um alelo preto, mas apenas em metade das células produzidas por meiose. (A metade remanescente não terá o alelo preto.)

(b) Nos porquinhos-da-índia marrons F_1 (*bb*), apenas os homólogos têm o alelo marrom, então o número de cópias presentes em cada estágio é: G_1, dois alelos marrons, G_2, quatro alelos marrons, metáfase da mitose, quatro alelos marrons; metáfase I da meiose, quatro alelos marrons; metáfase II, dois alelos marrons, após a citocinese da meiose, um alelo marrom.

33. **(a)** 9/16 preto e curvo, 3/16 preto e normal, 3/16 cinza e curvo e 1/16 cinza e normal.

(b) 1/4 preto e curvo, 1/4 preto e normal, 1/4 cinza e curvo e 1/4 cinza e normal.

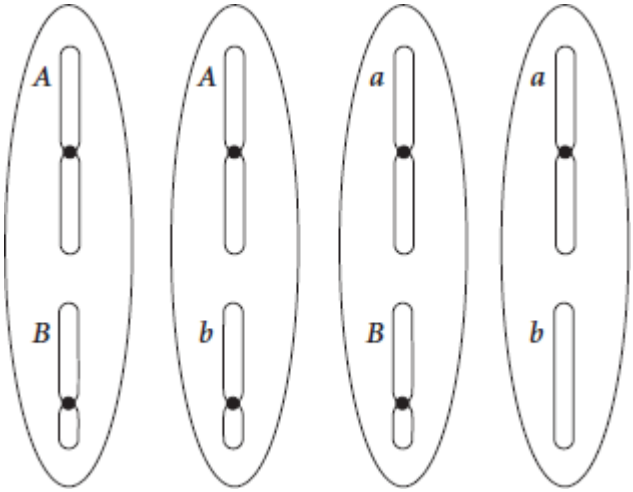
34. **(a)** 1/2 (*Aa*) × 1/2 (*Bb*) × 1/2 (*Cc*) × 1/2 (*Dd*) × 1/2 (*Ee*) = 1/32

(b) 1/2 (*Aa*) × 1/2 (*bb*) × 1/2 (*Cc*) × 1/2 (*dd*) × 1/4 (*ee*) = 1/64

(c) 1/4 (*aa*) × 1/2 (*bb*) × 1/4 (*cc*) × 1/2 (*dd*) × 1/4 (*ee*) = 1/256

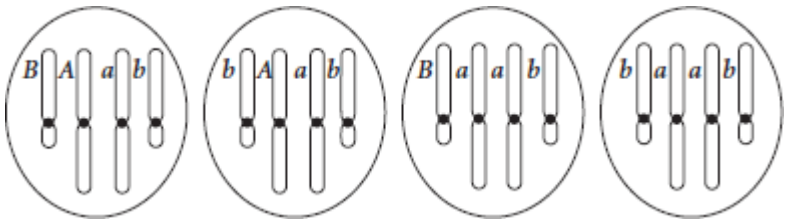
(d) Nenhum descendente tem esse genótipo. O genitor *Aa Bb Cc dd Ee* não pode contribuir com o alelo *D*; o genitor *Aa bb Cc Dd Ee* não pode contribuir com o alelo *B*. Portanto, seus descendentes não podem ser homozigotos para os *loci* dos genes *BB* e *DD*.

37. **(a)** Gametas do indivíduo *Aa Bb*:

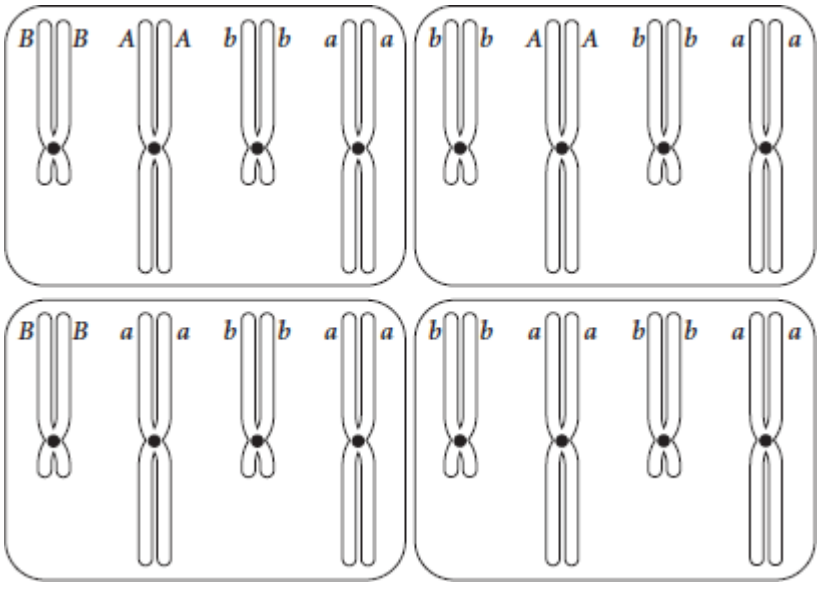


Gametas do indivíduo *aa bb*:

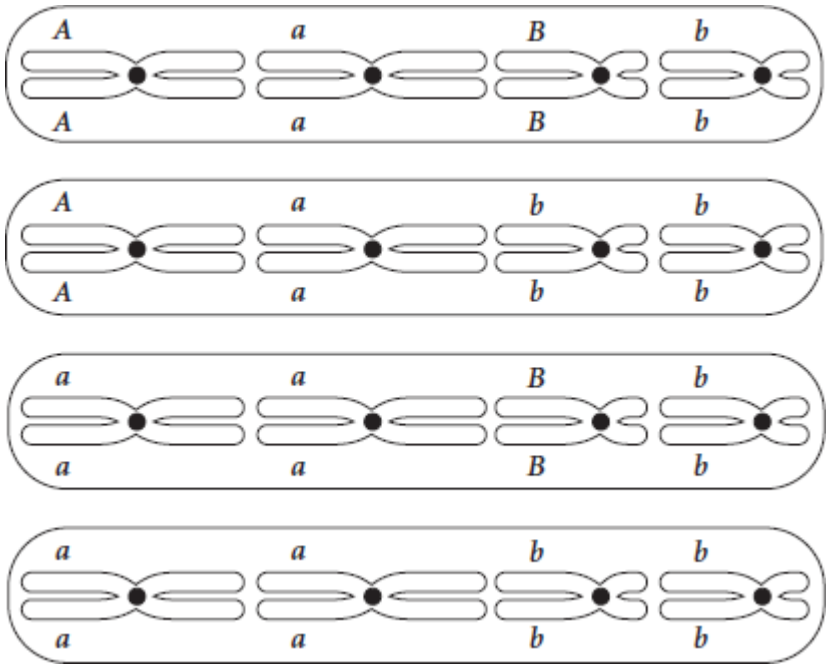
(b) Descendentes em G_1:



Descendentes em G_2:



Descendentes na metáfase da mitose:



A ordem dos cromossomos na placa metafásica pode variar.

38. **(a)** O cruzamento burnsi × burnsi produziu descendentes burnsi e pipiens, sugerindo que os genitores eram heterozigotos com cada um tendo um alelo burnsi e um alelo pipiens. O cruzamento também sugere que o alelo burnsi é dominante sobre o alelo pipiens. Os descendentes dos cruzamentos burnsi × pipiens sugerem que cada um dos cruzamentos era entre um sapo homozigoto recessivo (pipiens) e um sapo heterozigoto dominante (burnsi). Os resultados de ambos os cruzamentos são consistentes com o fenótipo burnsi sendo recessivo em relação ao fenótipo pipiens.

    (b) *B* representa o alelo burnsi e B^+ representa o alelo pipiens.

    Burnsi (BB^+) × burnsi (BB^+)

    Burnsi (BB^+) × pipiens (B^+B^+)

    Burnsi (BB^+) × pipiens (B^+B^+)

    (c) Para o cruzamento burnsi × burnsi ($BB^+ \times BB^+$), esperaríamos uma razão fenotípica de 3:1 nos descendentes. Um teste qui-quadrado para avaliar a adequação dos números observados de descendentes com uma razão esperada de 3:1 fornece um valor de qui-quadrado de 3,26 com 1 grau de liberdade. A probabilidade associada com este valor de qui-quadrado está entre 0,1 e 0,05, indicando que as diferenças entre o que esperamos e o que observamos poderiam ter surgido ao acaso.

    Para o primeiro cruzamento burnsi × pipiens ($BB^+ \times B^+B^+$), esperaríamos uma razão fenotípica de 1:1. Um teste qui-quadrado comparando os valores observados e esperados gera $\chi^2 = 1{,}78$, gl = 1, $P > 0{,}05$. Para o segundo cruzamento burnsi × pipiens ($BB^+ \times B^+B^+$), esperaríamos uma razão fenotípica de 1:1. Um teste qui-quadrado de adequação dos números observados com os esperados com uma razão 1:1 gera $\chi^2 = 0{,}46$, gl= 1, $P > 0{,}05$. Assim, todos os três cruzamentos são consistentes com a previsão de que o alelo burnsi é dominante sobre o alelo pipiens.

41. **(a)** Para o cruzamento de uma planta F_1 heterozigota (*Cc Ff*) e uma planta homozigota recessiva (*cc ff*) esperaríamos uma razão fenotípica de 1:1:1:1 para as diferentes classes fenotípicas. Um teste qui-quadrado de adaptação dos números observados com os esperados com uma razão 1:1:1:1 gera um valor de qui-quadrado 35, com gl = 3, $P < 0{,}005$.

    (b) A partir do valor de qui-quadrado, é improvável que o acaso produza as diferenças entre as razões observadas e esperadas, indicando que os descendentes não estão em uma razão 1:1:1:1.

    (c) O número de plantas com genótipo *cc ff* é menor que o esperado. O genótipo *cc ff* é possivelmente subletal (as papoulas-da-califórnia com os genótipos homozigotos recessivos são possivelmente menos viáveis que os outros genótipos possíveis).

43. O primeiro geneticista que identificou um alelo para obesidade acreditava que o alelo era recessivo. Vamos chamar seu alelo para obesidade o_1 e o alelo normal O_1. Com base nos cruzamentos que o geneticista fez, parecia que o alelo para obesidade era recessivo.

Cruzamento 1 com possível genótipo:

Obeso (o_1o_1) × Normal (O_1O_1)

↓

F_1 Todos normais (O_1o_1)

Cruzamento 2 com possíveis genótipos:

F_1 Normal (O_1o_1) × Normal (O_1o_1)

↓

F_2 8 normais (O_1O_1 e O_1o_1)
2 obesos (o_1o_1)

Cruzamento 3 com possíveis genótipos:

Obeso (o_1o_1) × Obeso (o_1o_1)

↓

F_1 Todos obesos (O_1o_1)

Vamos chamar seu segundo alelo do geneticista para obesidade o_2 e o alelo normal O_2. O cruzamento entre os camundongos obesos de dois laboratórios produziu apenas camundongos normais. Os alelos para obesidade de ambos os laboratórios eram recessivos. Entretanto, estavam localizados em diferentes *loci* de genes. Em resumo, os camundongos obesos de diferentes laboratórios têm genes separados para obesidade que são independentes um do outro.

Os prováveis genótipos para os camundongos obesos são:

Camundongo 1 obeso $(o_1o_1\ O_2O_2)$ × Camundongo 2 obeso $(O_1O_1\ o_2o_2)$

↓

F_1 Todos normais $(O_1o_1\ O_2o_2)$

Capítulo 4

4. Sistema XX-XY; os machos são heterogaméticos e produzem gametas com um cromossomo X ou um cromossomo Y. No sistema ZZ-ZW; as fêmeas são heterogaméticas e produzem gametas com um cromossomo Z ou W.

14. Não, se cada par XY for separado de forma independente, os gametas teriam números diferentes de cromossomos X e Y, resultando em descendentes com muitas combinações diferentes dos cromossomos X e Y, por exemplo X1X1Y2Y2X3Y3X4X4Y5Y5.

15. **(a)** Fêmea; **(b)** macho; **(c)** macho, estéril; **(d)** fêmea; **(e)** macho; **(f)** fêmea; **(g)** metafêmea; **(h)** macho, **(i)** intersexo; **(j)** fêmea; **(k)** metamacho, estéril; **(l)** metamacho; **(m)** intersexo.

17. **(a)** Fêmea; **(b)** macho; **(c)** macho; **(d)** fêmea; **(e)** macho.

21. A mãe do seu pai e o pai de seu pai; mas não a mãe da sua mãe ou o pai de sua mãe.

22. As excepcionais fêmeas de olhos brancos de Bridges eram X^wX^wY e os machos de olhos vermelhos eram X^+Y. Os resultados do cruzamento de tais fêmeas de olhos brancos e os machos de olhos vermelhos estão apresentados na Figura 4.11. A meiose nesta fêmea gera 45% de gametas X^wY, 45% de gametas X^w, 5% de gametas X^wX^w e 5% de gametas Y. A meiose no macho de olhos vermelhos gera 50% de gametas X^+ e 50% de gametas Y.

	0,5 X^+	**0,5 Y**
0,45 X^wY	0,225 X^wX^+Y fêmeas de olhos vermelhos	0,225 X^wYY machos de olhos brancos
0,45 X^w	0,225 X^wX^+ fêmeas de olhos vermelhos	0,225 X^wY machos de olhos brancos
0,05 X^wX^w	0,025 $X^wX^wX^+$ metafêmea, morre	0,025 X^wX^wY fêmeas de olhos brancos
0,05 Y	0,025 X^+Y machos de olhos vermelhos	0,025 YY Morre

23. **(a)** Sim; **(b)** sim; **(c)** não; **(d)** não.

24. **(a)** F_1: 1/2 X^+Y (machos cinza), 1/2 X^+X^y (fêmeas cinza); F_2: 1/4 X^+Y (machos cinza), 1/4 X^yY (machos amarelos), 1/4 X^+X^y (fêmeas cinza), 1/4 X^+X^+ (fêmeas cinza).

(b) F_1: 1/2 X^yY (machos amarelos), 1/2 X^+X^y (fêmeas cinza); F_2: 1/4 X^+Y (machos cinza), 1/4 X^yY (machos amarelos), 1/4 X^+X^y (fêmeas cinza), 1/4 X^yX^y (fêmeas amarelas).

(c) F_2: 1/4 X^+Y (machos cinza), 1/4 X^yY (machos amarelos), 1/4 X^+X^+ (fêmeas cinza), 1/4 X^+X^y (fêmeas cinza).

(d) F_3: 1/8 machos cinza, 3/8 machos amarelos, 5/16 fêmeas cinza e 3/16 fêmeas amarelas.

26. Como o daltonismo é um traço recessivo, a filha daltônica tem de ser homozigota recessiva. Como o daltonismo está ligado ao X, há motivos para John ter suspeitas. Normalmente, sua filha herdaria o cromossomo X de John. Como John não é daltônico, ele não poderia ter transmitido um cromossomo X com o alelo para daltonismo para sua filha.

Uma possibilidade remota é que sua filha seja XO, tendo herdado um alelo recessivo para daltonismo de sua mãe e nenhum cromossomo sexual de seu pai. Neste caso, ela teria a síndrome de Turner. Também é possível uma nova mutação do daltonismo ligada ao X, embora seja menos provável.

Se Cathy der à luz um filho daltônico, então John não tem motivos para suspeitar. O filho teria herdado o cromossomo Y de John e o cromossomo X com daltonismo de Cathy.

28. Como Bob precisa ter herdado o cromossomo Y de seu pai e seu pai tinha visão normal, um evento de não disjunção na linhagem paterna não pode ser responsável pelo genótipo de Bob. A mãe de Bob tem de ser heterozigota X^+X^c porque ela tem visão normal, e necessariamente herdou o cromossomo X com daltonismo do pai, daltônico. Para Bob herdar dois cromossomos X de daltonismo de sua mãe, o ovócito teria de ser resultante de uma não disjunção na meiose II. Na meiose I, os cromossomos X homólogos se separam e uma célula tem o cromossomo X^+ e outra tem o cromossomo X^c. A não separação das cromátides-irmãs na meiose II geraria um ovócito com duas cópias de X^c.

33. F_1: 1/2 Z^bZ^+ (machos normais); 1/2 Z^bW (fêmeas carecas). F_2: 1/4 Z^+Z^b (machos normais), 1/4 Z^+W (fêmeas normais), 1/4 Z^bZ^b (machos carecas), 1/4 Z^bW (fêmeas carecas).

38. **(a)** F_1: todos os machos têm asas pequenas e olhos vermelhos (X^mY s^+s) e todas as fêmeas têm asas longas e olhos vermelhos (X^+X^m s^+s). F_2: 3/16 macho, normais, vermelhos; 1/16 machos, normais, marrom-claros; 3/16 machos, miniaturas, vermelhos; 1/16 machos, miniaturas, marrom-claros; 3/16, fêmeas, normais, vermelhos, 1/16 fêmeas, normais, marrom-claros; 3/16 fêmeas, miniaturas, vermelhos; 1/16 fêmeas, miniaturas, marrom-claros.

(b) F_1: todas as fêmeas têm asas longas e olhos vermelhos (X^+X^m s^+s) e todos machos têm asas longas e olhos vermelhos (X^+ Y s^+s). F_2: 3/16 machos, asas longas, olhos vermelhos; 1/16 machos, asas longas, olhos marrom-claros; 3/16 machos, asas pequenas, olhos vermelhos; 1/16 machos, asas pequenas, olhos marrom-claros; 6/16 fêmeas, asas longas, olhos vermelhos; 2/16 fêmeas, asas longas, olhos marrom-claros.

40. A explicação trivial para estas observações é que esta forma de daltonismo é um traço autossômico recessivo. Neste caso, o pai seria um heterozigoto e esperaríamos proporções iguais de crianças daltônicas e normais de cada sexo.

Se, por outro lado, assumirmos que esta forma de daltonismo é um traço ligado ao X, a mãe é X^cX^c e o pai deve ser X^+Y. Normalmente, todos os filhos seriam daltônicos e todas as filhas devem ter visão normal. A forma mais provável de uma filha ser daltônica seria ela não ter herdado um X^+ do seu pai. A observação de que a filha daltônica tem baixa altura e não chegou à puberdade é consistente com a síndrome de Turner (XO). A filha daltônica então seria X^cO.

41. **(a)** 1; **(b)** 0; **(c)** 0; **(d)** 1; **(e)** 1; **(f)** 2; **(g)** 0; **(h)** 2; **(i)** 3.

44. **(a)** A inativação do X ocorre de forma aleatória em cada uma das células do embrião inicial e então esta inativação X é mantida nas células mitóticas dos descendentes. As manchas irregulares da pele sem glândulas sudoríparas surgem a partir de células precursoras da pele que inativam o cromossomo X com o alelo normal.

(b) O caso de inativação do X ocorre aleatoriamente em cada uma das células do embrião inicial. Mesmo em gêmeos idênticos, diferentes células precursoras ectodérmicas vão inativar diferentes cromossomos X, resultando em diferentes distribuições de manchas sem glândulas sudoríparas.

45. Partindo do pressuposto que o camundongo macho com testículos aumentados é fértil, vamos cruzá-lo com uma fêmea normal. Se o traço for autossômico dominante e o genitor for heterozigoto, apenas metade dos descendentes machos terá testículos aumentados. Se o traço for ligado ao X, todos os descendentes machos terão testículos aumentados.

Se todos os descendentes machos expressarem o traço, o traço ainda será autossômico dominante se o genitor for homozigoto. Se apenas alguns dos descendentes machos expressam o traço, a possibilidade ainda permanece de que o traço é ligado ao X, mas incompletamente penetrante. Em cada caso, serão necessárias mais evidências conclusivas.

Cruzar descendentes fêmeas (fêmeas F_1) com machos normais. Se o traço for autossômico dominante, alguns dos descendentes machos F_2 terão testículos aumentados. Se o traço for ligado ao Y, todos os descendentes machos F_2 terão testículos normais.

Capítulo 5

13. **(a)** Os resultados dos cruzamentos indicam que cremello e castanho são traços de linhagem pura (homozigotos). Palomino é um traço heterozigoto. A hipótese mais simples consistente com estes resultados é a dominância incompleta, com o palomino como o fenótipo dos heterozigotos resultantes dos castanhos cruzados com cremellos.

    (b) Considere C^B = castanho, C^W = cremello. Os genitores e descendentes destes cruzamentos têm os seguintes genótipos: castanho = C^BC^B; cremello = C^WC^W; palomino = C^BC^W.

15. Para ter longos lóbulos de orelha, a criança precisa herdar o alelo dominante e também expressá-lo. A probabilidade de herdar o alelo dominante é 50%; a probabilidade de expressá-lo é 30%. A probabilidade combinada de ambos é 0,5 × 0,3 = 0,15 ou 15%.

17. **(a)** A razão 2:1 nos descendentes de dois hamsters manchados sugere letalidade e a razão 1:1 nos descendentes de um hamster manchado e um hamster sem manchas indica que o manchado é um fenótipo heterozigoto. Se *S* e *s* representam alelos no *locus* para a mancha branca, os hamsters manchados são *Ss* e os hamsters de cor uniforme são *ss*. Um quarto dos zigotos esperados de um cruzamento de dois hamsters manchados são *SS* – embrionário letal – e ausentes nos descendentes, resultando na razão 2:1 de descendentes de manchas para cor uniforme.

    (b) Com a mancha é um fenótipo heterozigoto, é impossível conseguir hamsters chineses que gerem filhotes com manchas.

19. **(a)** Esperamos 1/4 Rp^1/Rp^1 (vermelho), 1/4 Rp^1/rp (vermelho), 1/4 Rp^2/Rp^1 (vermelho), 1/4 Rp^2/rp (púrpura), para toda razão fenotípica de 3/4 vermelho, 1/4 púrpura; **(b)** 1/4 Rp^1/Rp^2 (vermelho), 1/4 Rp^1/rp (vermelho), 1/4 Rp^2/rp (púrpura), 1/4 rp/rp (verde), para toda razão fenotípica de 1/2 vermelho, 1/4 púrpura, 1/4 verde; **(c)** é um cruzamento entre dois heterozigotos, então a razão fenotípica esperada é 3/4 vermelho, 1/4 púrpura; **(d)** é um cruzamento de um heterozigoto com um homozigoto recessivo. A razão fenotípica esperada é 1/2 púrpura, 1/2 verde. **(e)** 1/2 Rp^1/rp (vermelho), 1/2 Rp^2/rp (púrpura).

24. O genótipo da criança tem um alelo para tipo sanguíneo B e um alelo para tipo sanguíneo N que poderia não ter vindo da mãe e tem de ser oriundo do pai. Portanto, o pai da criança precisa ter um alelo para o tipo B e um alelo para tipo N. George, Claude e Henry são eliminados como possíveis pais porque eles não têm um alelo para o tipo B ou tipo N.

25. **(a)** Todos noz (*Rr Pp*); **(b)** 1/4 noz (*Rr Pp*), 1/4 rosa (*Rr pp*), 1/4 ervilha (*rr Pp*), 1/4 simples (*rr pp*); **(c)** 9/16 noz (*R_P_*), 3/16 rosa (*R_pp*), 3/16 ervilha (*rr P_*), 1/16 simples (*rr pp*); **(d)** 3/4 rosa (*R_ pp*), 1/4 simples (*rr pp*); **(e)** 1/4 noz (*Rr Pp*), 1/4 rosa (*Rr pp*), 1/4 ervilha (*rr Pp*), 1/4 simples (*rr pp*); **(f)** 1/2 rosa (*Rr pp*), 1/2 simples (*rr pp*).

26. **(a)** Cada um dos descendentes do retrocruzamento receberam alelos recessivos *b* e *r* do genitor branco homozigoto. Seu fenótipo é, portanto, determinado pelos alelos recebidos pelo genitor F_1: peixes marrons são *Bb Rr*; peixes azuis são *Bb rr*; peixes vermelhos são *bb Rr*, e peixes brancos são *bb rr*.

    (b) Esperamos uma razão 1:1:1:1 dos quatro fenótipos.

	Observado	**Esperado**	**O – E**	**$(O - E)^2/E$**
Marrom	228	229,25	–1,25	0,007
Azul	230	229,25	0,75	0,002
Vermelho	237	229,25	7,75	0,262
Branca	222	229,25	–7,25	0,229
Total	917	917		0,5 = χ^2

df (gl) = 4 – 1 = 3; 0,9 < P < 0,975

A probabilidade é alta de que a diferença entre o observado e o esperado seja por causa do acaso, e, portanto, aceitamos a razão 1:1:1:1.

(c) O peixe vermelho homozigoto seria *bb RR*, cruzado com *bb rr*. Todos os descendentes seriam *bb Rr* ou peixe vermelho.

(d) Peixe vermelho homozigoto *bb RR* × peixe azul homozigoto *BB rr*

Retrocruzamento: *Bb RR* (F_1) para *bb RR* (genitor vermelho homozigoto).

Descendentes do retrocruzamento serão: 1/4 *Bb RR* (marrom); 1/4 *Bb Rr* (marrom); 1/4 *bb RR* (vermelho); e 1/4 *bb Rr* (vermelho). No geral, 1/2 marrom e 1/2 vermelho.

29. **(a)** Uma vez que todos os filhotes são pretos, todos necessariamente herdaram um alelo *B* dominante do genitor amarelo e um alelo *E* dominante do genitor marrom. A genitora marrom tem de ser *bb EE* e o macho amarelo tem de ser *BB ee*. Todos os filhotes pretos eram *Bb Ee*.

    (b) Cruzar dois labradores amarelos produzirá apenas filhotes amarelos. Cruzar dois labradores marrons produzirá apenas filhotes marrons, se pelo menos um dos genitores for *EE* homozigoto ou 3/4 marrom e 1/4 amarelo se ambos os genitores forem *Ee* heterozigotos.

31. Considere que *A* e *B* representam dois *loci*. Os F_1 heterozigotos são *Aa Bb*. Os F_2 são: *A_ B_* formato de disco, *A_ bb* esférico, *aa B_* esférico, *aa bb* longo.

33. **(a)** De acordo com as informações no Quadro 5.3, os setters irlandeses são *BB ee SS* e *A* ou a^t. *B* permite a expressão do pigmento preto, mas o genótipo *ee* inibe a cor preta no pelo, resultando em uma cor avermelhada exceto no nariz e nos olhos. *S* inibe as manchas, resultando em uma cor uniforme.

    (b) Poodles são *SS*. Como o alelo *S* dominante inibe as manchas, nenhum filhote a partir dos cruzamentos com poodles terá manchas.

    (c) Os são-bernardos são $a^y a^y$ *BB* e os dobermanns são $a^t a^t$ *EE SS*. Os descendentes terão o genótipo $a^y a^t$ *B_E_S_*. Como a^y que especifica amarelo é dominante sobre a^t e o alelo *E* permite a expressão do genótipo *A*, os descendentes terão pelos amarelos.

    (d) Rottweilers são $a^t a^t$*BB EE SS* e os labradores retriever são $A^S A^S$ *SS*. Os descendentes serão $A^s a^t$ *B_E_SS*. A combinação dos alelos A^s e *E* criará cores uniformes de pelo.

35. Considere que H^+ represente o alelo para a ocorrência de chifres e H^- represente o alelo sem chifre. Como H^+ é recessivo nas fêmeas, a genitora com chifre tem de ser H^+H^+. O macho sem chifre é H^-H^- por causa da ausência de chifres nos machos recessivos. Seus descendentes F_1 serão todos H^+H^-. Um intercruzamento de F_1 produziria descendentes machos e fêmeas na razão de 1 H^+H^+, 2 H^+H^-, e 1 H^-H^-. Novamente, lembrando que H^+ é recessivo nas fêmeas, esperaríamos uma razão de 3:1 sem chifres para as fêmeas com chifre.

39. No efeito genético materno, o genótipo da genitora determina o fenótipo dos descendentes. Como Marta é canhota, sabemos que o genótipo de sua mãe tem de ser *ss*. Se a mãe da Marta for *ss*, Marta é carreadora obrigatoriamente de pelo menos um alelo *s*. Não temos informações sobre o pai da Marta. **(a)** Falso. Marta poderia ter herdado um s^+ do seu pai e, portanto, poderia ser s^+s; **(b)** verdadeiro. Marta precisa ter herdado um alelo *s* de sua mãe e, portanto, não pode ser s^+s^+; **(c)** falso. O fenótipo dos filhos de Marta será determinado pelo genótipo de Marta, que não conhecemos. Marta pode ter herdado um alelo s^+ de seu pai, no caso seu genótipo seria s^+s e ela teria apenas filhos destros; **(d)** falso. O genótipo de Marta poderia ser s^+s, neste caso todos seus filhos seriam canhotos; **(e)** falso. O fenótipo de sua mãe é determinado pelo seu próprio genótipo, que poderia ser s^+s; **(f)** verdadeiro. Como Marta é canhota, o genótipo de sua mãe é *ss* e todos seus filhos serão canhotos.

45. Responsável pelas fenocópias, temos 50% (subtraindo 1/3 de que são fenocópias de 75%) dos filhotes com o genótipo autossômico dominante para orelhas longas e 50% tendo o genótipo recessivo. Portanto, um genitor é homozigoto recessivo e o outro é um heterozigoto.

Capítulo 6

10. **(1)** A pessoa pode estar ciente de uma doença genética ou o fator de risco na sua família; **(2)** uma mulher mais idosa pode estar grávida ou contemplando a gravidez e pode precisar de informações sobre as opções para exame genético pré-natal; **(3)** uma pessoa pode ter teste positivo para uma doença genética ou fator de risco; **(4)** uma pessoa ou casal pode ter uma criança com uma doença genética e precisa de informação sobre seu risco de ter outros filhos com esta condição; **(5)** um casal pode ser parente (p. ex., primos de primeiro grau); **(6)** um casal tem dificuldade para engravidar ou manter a gravidez; **(7)** uma pessoa foi exposta a mutágenos ou substâncias químicas que provocam um risco maior de defeitos congênitos; **(8)** ambos os parceiros podem ser carreadores de uma condição genética recessiva; **(9)** um casal precisa de orientação na interpretação dos resultados de um teste pré-natal.

12. Na amniocentese é coletada uma amostra de líquido amniótico ao inserir uma agulha no saco amniótico, em geral feito nas 16 semanas de gravidez e requer a cultura das células fetais. A amostragem das vilosidades coriônicas pode ser feita na 10ª à 11ª semana de gravidez e consiste em um pequeno pedaço do cório por meio de inserção de um cateter na

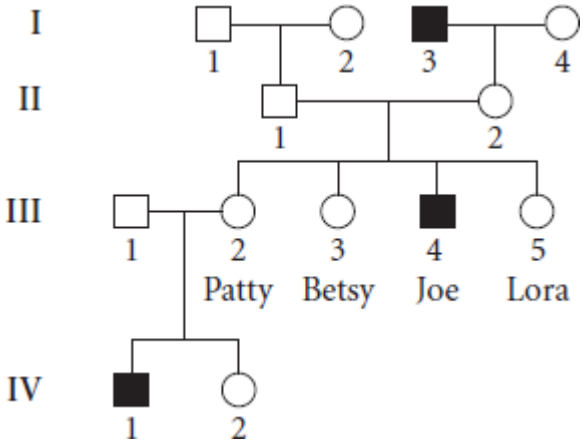
vagina. A amniocentese é relativamente segura, mas os resultados não estão disponíveis antes de 17 a 18 semanas de gravidez. A coleta de amostras das vilosidades coriônicas tem risco discretamente maior de complicação, incluindo lesão fetal, mas os resultados estão disponíveis algumas semanas antes.

15. Esses são os exames genéticos pedidos por leigos sem precisar de consentimento ou um pedido médico. Eles podem fazer testes para condições ou predisposições genéticas, paternidade ou ascendência. Existe uma preocupação de que os testes não sejam acurados ou de que a pessoa que os solicitar não esteja preparada para interpretar ou compreender adequadamente os resultados desses exames. Também existe uma preocupação sobre privacidade ou confidencialidade da pessoa que solicita o teste.

16. Esta lei norte-americana proíbe a discriminação por empregadores e seguradoras de saúde com base em informações obtidas por exame genético e impede estas empresas de pedir a uma pessoa que faça o exame genético para um emprego ou assistência de saúde.

18. O estudo da genética humana é necessário para compreender e superar as doenças genéticas humanas. Por causa do longo tempo de vida dos humanos, tamanho corporal relativamente grande e estilo de vida e comportamentos únicos, não existem modelos animais ou estes são insuficientes para muitos distúrbios genéticos. A preservação cuidadosa de registros de casamento, nascimento, morte e prontuários em muitas sociedades fornece uma riqueza de dados para análise genética. O término do projeto do genoma humano agora facilita mapear e identificar os genes humanos. Os humanos têm um forte senso de identidade e importância como indivíduos, e desejam compreender como o perfil genético de uma pessoa contribui para sua saúde, comportamentos, capacidades e incapacidades e chances futuras. O estudo da genética humana também pode revelar origens históricas e relações antropológicas de indivíduos e populações.

19. **(a)**



(b) Recessivo ligado ao X; **(c)** zero; **(d)** 1/4; **(e)** 1/4.

22. **(a)** Apenas os machos têm a doença, ela pula gerações e as fêmeas carreadoras não afetadas têm filhos afetados e não afetados. Estas observações estão consistentes com um traço recessivo ligado ao X. Os traços ligados ao Y são transmitidos diretamente do pai para o filho, não pulam gerações; **(b)** usaremos X^+ para indicar o alelo X normal e X^d para indicar o alelo Dent.

I: 1 – X^+Y; 2 – X^+X^d.

II: 1 e 5 são X^dY, 7 e 8 são X^+X^d e o resto não tem o alelo da doença.

III: 2 e 3 são carreadores X^+X^d, 4 a 7 são X^dY e o restante não tem o alelo da doença.

IV: 2 é X^dY, 1 é X^+Y.

24. **(a)** Dominante autossômico. O traço tem de ser autossômico porque os homens afetados transmitem o traço para filhos e filhas. É dominante porque não pula gerações, todos os indivíduos afetados têm pais afetados e é muito improvável que múltiplos indivíduos sem parentesco acasalado no heredograma sejam carreadores de um traço raro.

(b) Dominante ligado ao X. Superficialmente, esse heredograma parece ser semelhante ao heredograma na parte *a* pelo fato de que machos e fêmeas são afetados e parece ser um traço dominante. No entanto, uma análise mais meticulosa revela que enquanto as fêmeas afetadas transmitem o traço para filhos e filhas, os homens afetados transmitem o traço apenas para suas filhas.

(c) Ligado ao Y. O traço afeta apenas os homens e é transmitido de pai para filho. Todos os filhos do homem afetado são

afetados.

(d) Recessivo ligado ao X ou autossômico dominante limitado pelo sexo. Como apenas os homens mostram o traço, ele poderia ser recessivo ligado ao X, ligado ao Y ou limitado pelo sexo. Podemos eliminar a ligação ao Y porque os homens afetados não transmitem o traço para seus filhos. A herança recessiva ligada ao X é consistente com o padrão de carreadoras não afetadas que têm filhos afetados e não afetados e homens afetados que têm carreadoras não afetadas, mas sem filhos afetados. A herança autossômica dominante limitada ao sexo também é consistente com as mulheres heterozigotas não afetadas que têm filhos heterozigotos afetados, filhos homozigotos recessivos não afetados e filhas heterozigotas não afetadas ou homozigotas recessivas. As duas possibilidades remanescentes de recessivo ligado ao X *versus* autossômico dominante limitado pelo sexo poderiam ser diferenciadas se tivéssemos dados suficientes para determinar se os homens afetados têm filhos afetados e não afetados, como seria esperado da herança autossômica dominante ou se os homens afetados têm apenas filhos não afetados, como seria esperado da herança recessiva ligada ao X. Infelizmente, este heredograma mostra apenas dois filhos dos homens afetados. Em ambos os casos, os filhos não são afetados, consistente com herança recessiva ligada ao X, mas dois descendentes do sexo masculino não são suficientes para concluir que os homens afetados têm filhos afetados.

(e) Recessivo autossômico. Os pais não afetados têm descendentes afetados, então o traço é recessivo. A filha afetada herdou obrigatoriamente os alelos recessivos dos genitores não afetados, então o traço tem de ser autossômico. Se ele era ligado ao X, seu pai mostraria o traço.

30. Enxaquecas: genético e ambiental. Concordância acentuadamente maior em gêmeos monozigóticos, que são 100% geneticamente idênticos, do que em gêmeos dizigóticos, que são 50% geneticamente idênticos, é indicativa de influência genética. Entretanto, apenas 60% de concordância para gêmeos monozigóticos indica que os fatores ambientais têm participação significativa.

    Cor dos olhos: genético. A concordância é maior em gêmeos monozigóticos do que nos dizigóticos. Além disso, os gêmeos monozigóticos têm 100% de concordância para este traço, indicando que o ambiente não tem influência detectável.

    Sarampo: sem influência genética detectável. Não existe diferença na concordância entre os gêmeos monozigóticos e dizigóticos. Alguma influência ambiental pode ser detectada porque os gêmeos monozigóticos mostram menos de 100% de concordância.

    Pé torto: genético e ambiental. A justificativa para enxaqueca serve para este caso. Uma forte influência ambiental é indicada pela alta discordância nos gêmeos monozigóticos.

    Pressão arterial elevada: genética e ambiental. A justificativa é semelhante à do pé torto.

    Lateralidade: sem influência genética. A concordância é a mesma nos gêmeos monozigóticos e dizigóticos. A influência ambiental é indicada como menos de 100% de concordância nos gêmeos monozigóticos.

    Tuberculose: sem influência genética. A concordância é a mesma nos gêmeos monozigóticos e dizigóticos. A importância da influência ambiental é indicada pela discordância muito baixa nos gêmeos monozigóticos.

33. Dados sugerem que a esquizofrenia tem um forte componente genético. É muito mais provável que os pais biológicos dos esquizofrênicos adotados apresentem esquizofrenia do que os indivíduos sem parentesco (pais adotivos), apesar do fato de que os esquizofrênicos adotados compartilham o mesmo ambiente que os pais adotivos. Se as variáveis ambientais (como substâncias químicas na água ou alimento ou linhas de energia) forem um fator importante, seria possível observar maior frequência de esquizofrenia nos pais adotivos. Outra possibilidade é que esta maior frequência de esquizofrenia nos pais biológicos simplesmente reflete a maior probabilidade de que pais esquizofrênicos entreguem seus filhos para adoção. Este caso é descartado pelos dados de que os pais biológicos de adotados não esquizofrênicos não mostram maior frequência semelhante de esquizofrenia em comparação com os pais adotivos.

Capítulo 7

13. **(a)** Observaríamos segregação independente entre os marcadores genéticos no cromossomo X e a calvície padrão. Os traços influenciados pelo sexo são codificados por genes autossômicos que se distribuem independentemente dos genes ligados ao X, pois estão em diferentes cromossomos.

    (b) Se a calvície padrão for recessiva ligada ao X, ela seria codificada por um gene presente no cromossomo X. Se este gene e os marcadores ligados ao X estiverem próximos do cromossomo X, eles não se distribuem independentemente. Entretanto, se o gene para a calvície padrão e os marcadores ligados ao X estiverem muito afastados, então o *crossing over* entre eles ocorre a cada meiose, eles podem se distribuir independentemente.

15. **(a)** 1/4 olhos do tipo selvagem, asas do tipo selvagem; 1/4 olhos vermelhos, asas do tipo selvagem; 1/4 olhos do tipo selvagem, asas listradas brancas; 1/4 olhos vermelhos, asas listradas brancas; **(b)** o F_1 heterozigoto herdou um cromossomo com alelos para olhos vermelhos e asas listradas brancas (*re wb*) de um genitor e um cromossomo com alelos para olhos do tipo selvagem e asas do tipo selvagem (re^+ wb^+) de outro genitor. Estes são os fenótipos de descendente não recombinante, presente em maiores números. Os recombinantes são 19 com olhos vermelhos, asas do tipo selvagem e 16 com olhos do tipo selvagem, asas listradas brancas. RF = recombinantes/descendentes totais × 100% = (19 + 16)/879 × 100% = 4%

    A distância entre os genes é 4 unidades de mapa.

16. Os genes estão ligados e não se distribuem independentemente.

17. Formato de coração, numerosos espinhos, 33,7%.

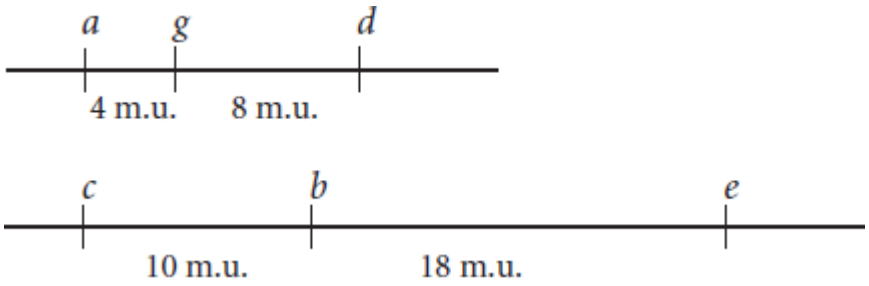
    Formato normal, poucos espinhos, 33,7%.

    Formato de coração, poucos espinhos, 16,3%.

    Formato normal, numerosos espinhos, 16,3%.

21. As distâncias entre os genes estão indicadas pelas taxas de recombinação. Como os *loci R* e L_2 têm a maior taxa de recombinação, eles devem estar muito distantes, e W_2 está no meio. A ordem dos genes é: *R*, W_2, L_2.

27.



    O gene *f* não está ligado a nenhum destes genes, ele está no terceiro grupo de ligação.

29. **(a)** Os não recombinantes têm *p sh-1* Hb^2 ou *P Sh-1* Hb^1; os *crossing overs* duplos têm *p Sh-1* Hb^2 ou *P sh-1* Hb^1. As duas classes diferem no *locus Sh-1*; portanto *Sh-1* é o *locus* do meio.

    (b) *P e Sh-1*: os recombinantes têm *P sh-1* ou *p Sh-1*. RF = (57 + 45 + 1)/708 = 0,145 = 14,5 m.u.

    Sh-1 e *Hb*: os recombinantes têm *Sh-1* Hb^2 ou *sh-1* Hb^1. RF = (6 + 5 + 1)/708 = 0,017 = 1,7 m.u.

    (c) *Crossing overs* duplos esperados = RF1 × RF2 × descendentes totais = 0,145(0,017)(708) = 1,7.

    Coeficiente de coincidência = *crossing overs* duplos observados/*crossing overs* duplos esperados = 1/1,7 = 0,59

    Interferência = 1 – coeficiente de coincidência = 0,41.

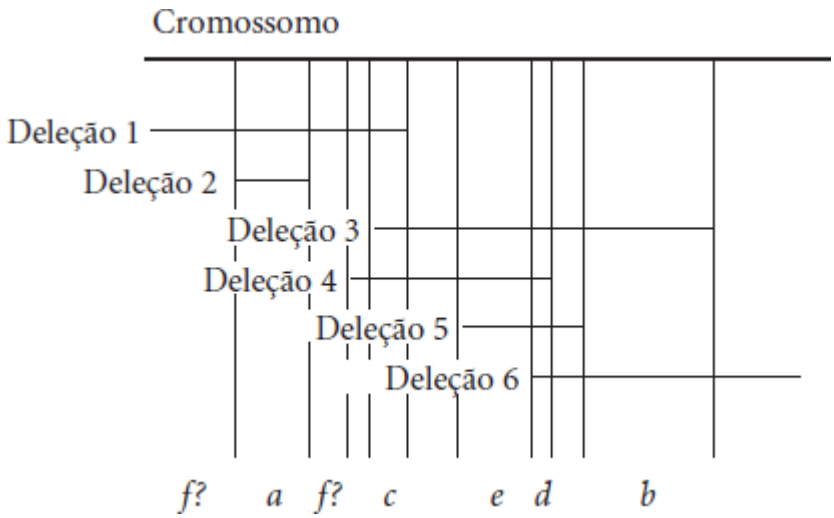
36. Este *lod score* indica que a probabilidade de observar este grau de associação se o marcador estiver ligado a asma é 100 vezes maior do que se o marcador não tiver ligação a asma. Um *lod score* de 3, para um risco relativo de 1.000 vezes, em geral é considerado uma evidência convincente de ligação.

38. A enzima 1 está localizada no cromossomo 9. O cromossomo 9 é o único cromossomo que está presente nas linhagens celulares que produzem a enzima 1 e ausente nas linhagens celulares que não produzem a enzima 1.

    A enzima 2 está localizada no cromossomo 4. O cromossomo 4 é o único que está presente nas linhagens celulares que produzem a enzima 2 (C e D) e ausente nas linhagens celulares que não produzem a enzima 2 (A e B).

    A enzima 3 está localizada no cromossomo X. O cromossomo X é o único cromossomo que está presente nas linhagens celulares que produzem a enzima 3 e ausente nas linhagens celulares que não produzem a enzima 3.

39.



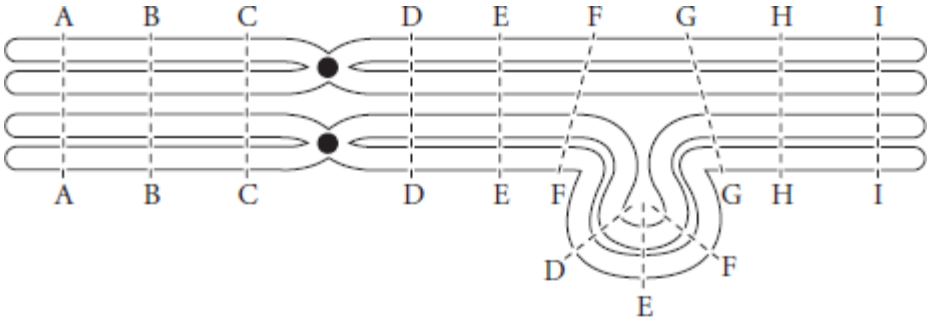
A localização de *f* é ambígua, ele poderia estar na localização apresentada no mapa de deleção.

Capítulo 8

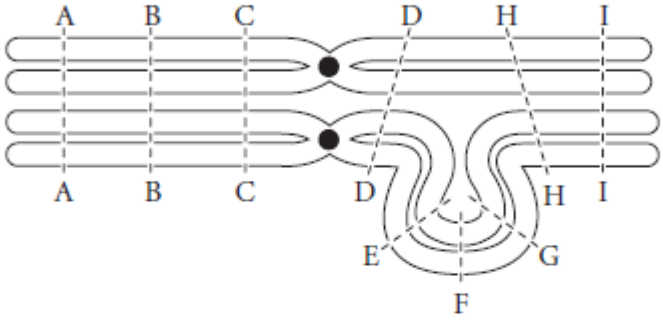
19. **(a)** Duplicações; **(b)** poliploidia; **(c)** deleções; **(d)** inversões; **(e)** translocações.
20. **(a)** Duplicação em *tandem* de AB; **(b)** duplicação deslocada de AB.

    (c) inversão paracêntrica de DEF; **(d)** deleção de B; **(e)** deleção de FG.

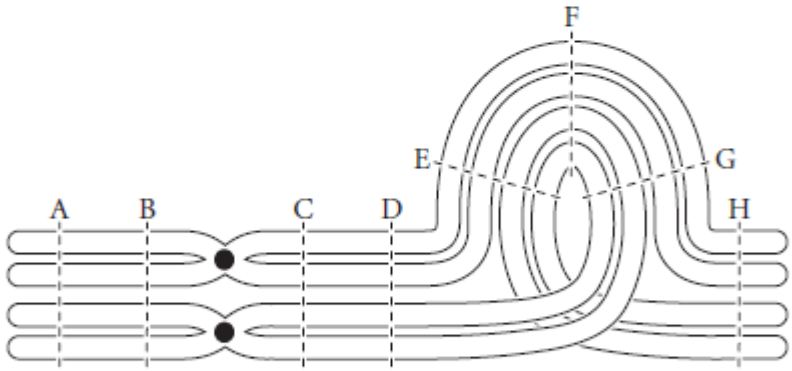
    (f) inversão paracêntrica de CDE; **(g)** inversão pericêntrica de ABC; **(h)** duplicação e inversão de DEF; **(i)** duplicação de CDEF; inversão de EF.
23. **(a)** 1/3 fêmeas de olhos brancos *Notch*; 1/3 de fêmeas do tipo selvagem e 1/3 machos do tipo selvagem; **(b)** 1/3 fêmeas de olhos vermelhos *Notch*; 1/3 fêmeas do tipo selvagem e 1/3 machos de olhos brancos; **(c)** 1/3 fêmeas de olhos brancos *Notch*; 1/3 fêmeas de olhos brancos e 1/3 machos de olhos brancos.
25. **(a)**



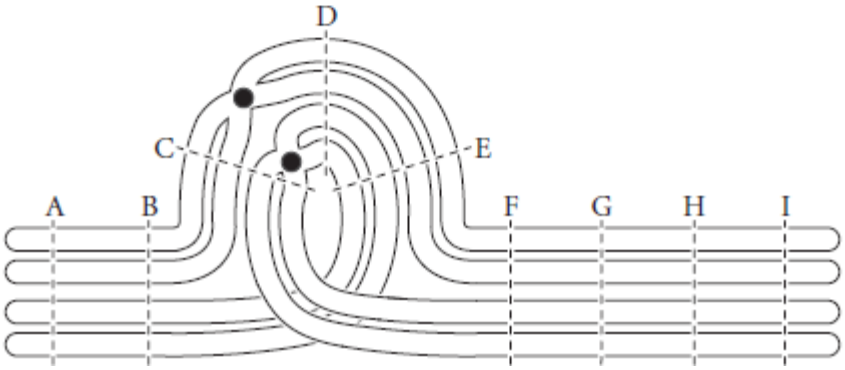
(b)



(c)

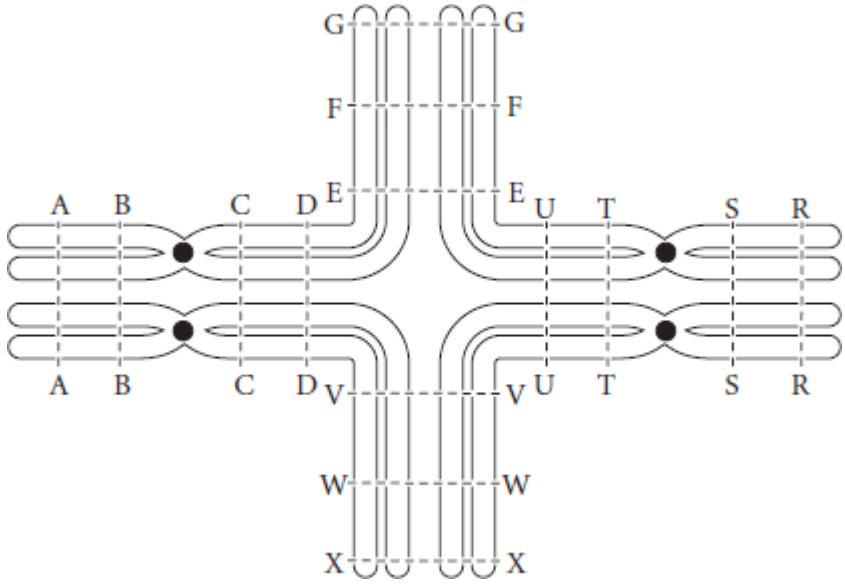


(d)

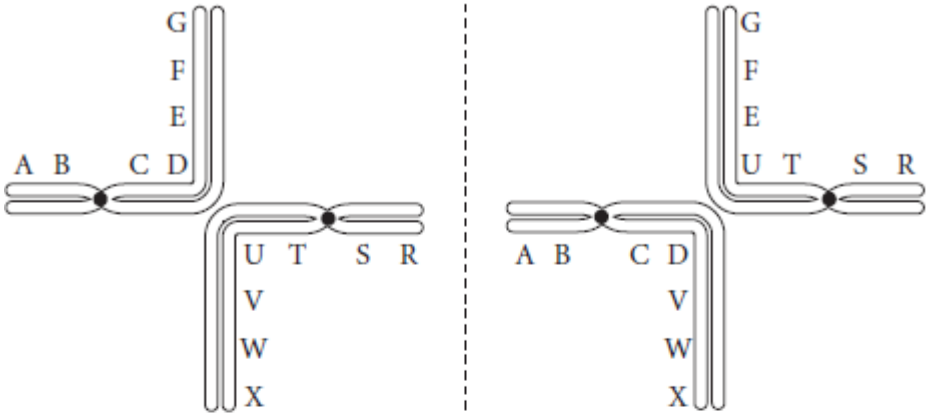


27. Cada evento de *crossing over* resulta em dois gametas recombinantes e dois não recombinantes. Se ocorre um *crossing over* em 100% das meioses, o resultado seria 50% de gametas recombinantes. Se ocorre *crossing over* dentro da inversão pericêntrica em uma taxa de 26% das meioses, então 13% dos ovócitos da mulher terão duplicação/cromossomo 8 deficiente. Se todos estes ovócitos foram ovos viáveis e se eles não resultarem em aborto precoce após a fertilização então a probabilidade de o casal ter uma criança com a síndrome causada pelo *crossing over* é 13%.

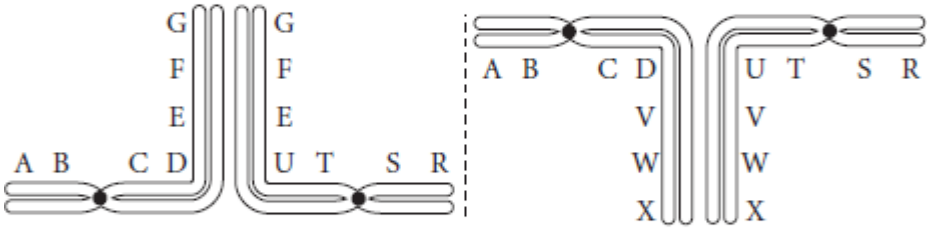
28. **(a)**



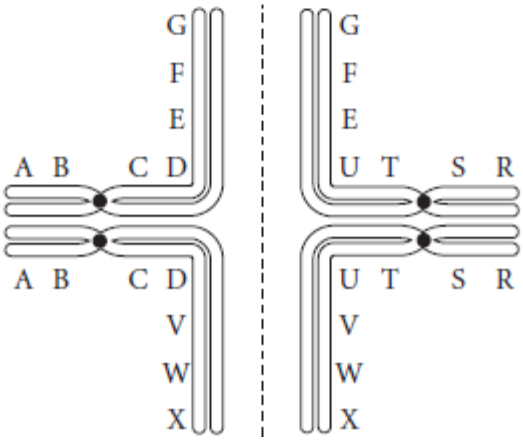
(b) Alternado:



Adjacente-1:



Adjacente-2:



(c) Alternado: os gametas têm cromossomos normais ou de translocação e são todos viáveis.

A B • C D E F G + R S • T U V W X

e

A B • C D V W X + R S • T U E F G

Adjacente-1: Os gametas têm um cromossomo normal e um cromossomo de translocação, resultando na duplicação de alguns genes e na deficiência de outros.

A B • C D E F G + R S • T U E F G

e

A B • C D V W X + R S • T U V W X

Adjacente-2 (raro): Os gametas têm um cromossomo normal e um cromossomo de translocação, com a duplicação de alguns genes e a deficiência de outros.

A B • C D E F G + A B • C D V W X

e

R S • T U V W X + R S • T U E F G

29. Como o pai tem visão normal, a mãe deve ser carreadora do daltonismo. O homem daltônico com síndrome de Klinefelter deve ter herdado duas cópias do cromossomo X com daltonismo de sua mãe. O evento de não disjunção ocorreu durante a meiose II do ovócito.

32. A alta incidência de síndrome de Down na família de Bill e entre os seus parentes é consistente com a síndrome de Down familiar, causada por uma translocação robertsoniana do cromossomo 21. Bill e sua irmã, que não estão afetados, são carreadores fenotipicamente normais da translocação e têm 45 cromossomos. Seus filhos e o irmão de Bill, que tem a síndrome de Down, têm 46 cromossomos. A partir das informações anteriores, não existe motivo para suspeitar que a esposa de Bill, Betty, tenha alguma anomalia cromossômica. Portanto, é mais provável que a afirmação na parte *d* esteja correta. Todas as outras afirmações estão erradas.

34. **(a)** Possíveis tipos de gametas: (i) cromossomo 13 normal e cromossomo 22 normal; (ii) cromossomos 13 + 22 translocados; (iii) cromossomos 13 + 22 translocados e cromossomo 22 normal; (iv) cromossomo 13 normal; (v) cromossomo 13 normal e cromossomos 13 + 22 translocados; (vi) cromossomo 22 normal.

(b) Tipos de zigotos: (i) 13, 13, 22, 22; normal; (ii) 13, 13 + 22, 22; carreador de translocação; (iii) 3, 13 + 22, 22, 22; trissomia 22; (iv) 13, 13, 22; monossomia 22; (v) 13, 13, 13 + 22, 22; trissomia 13; (vi) 13, 22, 22; monossomia 13.

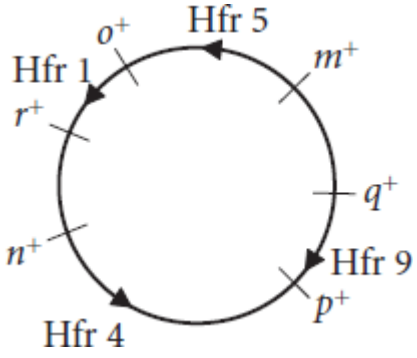
(c) 50%.

38. **(a)** Tais alotriploides podem ter 1*n* da espécie I e 2*n* da espécie II para 3*n* = 19; por outro lado, eles podem ter 2*n* da espécie I e 1*n* da espécie II para 3*n* = 15; **(b)** 4*n* = 18; **(c)** 2*n* + 1 = 9; **(d)** 2*n* – 1 = 13; **(e)** 2*n* +2 = 10; **(f)** Alotetraploides devem ter cromossomos de ambas espécies, e um total de 4*n*. Existem três combinações possíveis para tais alotetraploides: 2*n* em cada um: 2(4) + 2(7) = 22; 1*n* da espécie I + 3*n* da espécie II: 1(4) + 3(7) = 25; 3*n* da espécie I + 1*n* da espécie II: 3(4) + 1(7) = 19.

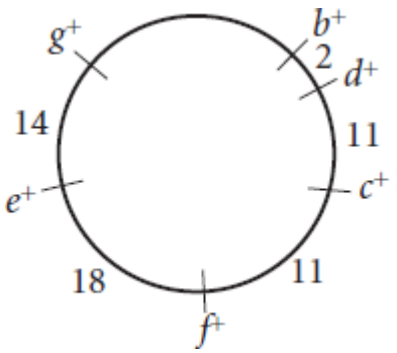
Capítulo 9

17. Isole o DNA de amostras de água retirada de dois rios. Sequencie o DNA e identifique as espécies de bactérias com base nas diferenças de seu DNA.

22. Em cada uma das cepas Hfr, o fator F foi inserido em uma diferente localização no cromossomo. A orientação do fator F nas cepas também varia.



23. As distâncias entre os genes estão em minutos.



25. Durante a transformação, apenas os genes que estão ligados perto ou localizados próximo um do outro no cromossomo doador serão transformados juntos. Em outras palavras, uma maior frequência de cotransformação indica uma distância mais curta entre os dois genes no cromossomo do doador.

    Para calcular a frequência de cotransformação dos genes trp^+ e tyr^+ do *Bacillus subtilis*, divida o número de células transformadas com o genótipo trp^+ tyr^+ pelo número total de células transformadas (354/820). A frequência da cotransformação é 0,43 ou 43%. O alto nível de cotransformação indica que estes genes estão intimamente ligados.

31. Primeiro, devemos identificar o fago descendente cujo fenótipo de placa é diferente dos fagos infectantes. Os fagos infectantes originais eram tipo selvagem (a^+ b^+) e duplamente mutante (*a b*). Qualquer fago que dê origem ao fenótipo de placa a^+ b^- ou ao fenótipo de placa a^- b^+ foram produzidos por recombinação entre os dois tipos de partículas de fago infectante.

Fenótipo da placa	Número
a^+ *b*	2.043
a^+ *b*	320 (recombinante)
a^- b^+	357 (recombinante)
a^- b^-	2.134
Total de placas	4.854

    A frequência da recombinação é calculada ao dividir o número total de placas recombinantes pelo número total de placas (677/4.854), que dá uma frequência de 0,14 ou 1 = 14%.

32. **(a)** Transdutantes foram inicialmente analisados para a presença de $proC^+$. Assim, apenas os transdudantes $proC^+$ foram identificados.

    (b) Os genótipos do tipo selvagem ($proC^+$ $proA^+$ $proB^+$ $proD^+$) representam transdutantes únicos de $proC^+$. Ambos os genótipos $proC^+$ $proA^-$ $proB^+$ $proD^+$ e $proC^+$ $proA^+$ $proB^-$ $proD^+$ representam cotransdutantes de $proC^+$, $proA^-$ e $proC^+$, $proB^-$.

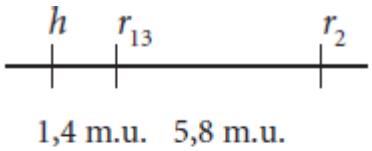
    (c) Os genótipos do tipo selvagem ($proC^+$ $proA^+$ $proB^+$ $proD^+$) representam transdutantes únicos de $proC^+$. Ambos os genótipos $proC^+$ $proA^-$ $proB^+$ $proD^+$ e $proC^+$ $proA^+$ $proB^-$ $proD^+$ representam cotransdutantes de $proC^+$, $proA^-$ e $proC^+$ $proB^-$.

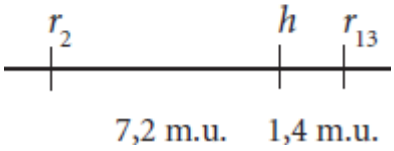
33. Sabemos que os dois genes estão a 8 unidades de mapa de distância. Essas 8 unidades de mapa correspondem à recombinação de porcentagem entre os dois genes de 8%. Quando a geneticista mistura dois fagos (m^+ c^+ × m^- c^-), criando uma infecção dupla da célula bacteriana, ela vai esperar os dois tipos de fenótipos de placa recombinantes, m^+ c^- *e* m^- c^+, para comprometer 8% dos fagos descendentes. Os 92% remanescentes serão uma combinação do fago do tipo selvagem e o fago duplamente mutante.

Fenótipo da placa	Número esperado
c^+ m^+	460
c^- m^-	460
c^+ m^-	40 (recombinante)
c^+ m^-	40 (recombinante)
Total de placas	1.000

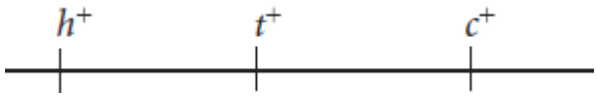
35. **(a)** A frequência de recombinação entre r_2 e *h* é13/180 = 0,072 ou 7,2%. O RF entre r_{13} e *h* é 13/216 = 0,014 ou 1,4%.

    (b)

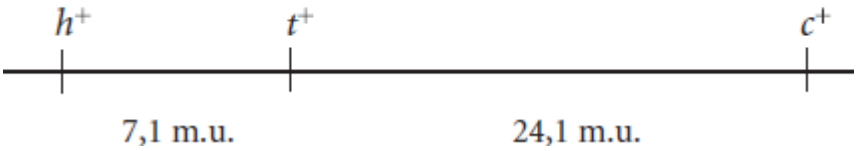




36. **(a)**



(b)



(c) Coeficiente de coincidência = (6 + 5)/(0,071 × 0,241 × 942) = 0,68. Interferência = 1 – 0,68 = 0,32.

40. **(a)** Se ocorre complementação entre os diferentes mutantes fago que são infectantes das células K da *E. coli*, então as placas formarão um tapete de células K da *E. coli*. A complementação pode ocorrer apenas quando as mutações de diferentes fagos estão localizadas em diferentes cístrons ou genes funcionais. Os fagos mutantes que não complementam um ao outro têm mutações que estão nos mesmos cístrons. Existem três grupos de fagos que não se complementam com outros fagos dentro do seu grupo, mas se complementam aos fagos nos outros grupos. Como existem três grupos, podemos deduzir a presença de três cístrons ou genes funcionais.

(b) Grupo 1: mutantes 1, 5, 6 e 8

Grupo 2: mutantes 2, 4 e 7

Grupo 3: mutante 3

Capítulo 10

18. Ele tirou imagens de difração por raios X usadas na construção da estrutura do DNA: **(i)** Wilkins e Franklin; determinaram que o DNA tem bases nitrogenadas. **(a)** Kossel; identificou o DNA como material genético no bacteriófago. **(f)** Hershey e Chase; descobriram a regularidade nas proporções de diferentes bases no DNA. **(j)** Chargaff; determinou que o DNA é responsável pela transformação nas bactérias. **(g)** Avery, MacLeod e McCArty; trabalharam na estrutura helicoidal do DNA ao construir modelos. **(c)** Watson e Crick; descobriram que o DNA é composto por nucleotídios repetidos. **(d)** Levene; determinou que o DNA é ácido e com alto teor de fósforo. **(e)** Miescher; conduziu experimentos mostrando que o RNA pode servir como material genético em alguns vírus. **(b)** Fraenkel-Conrat; demonstram que o material morto por calor das bactérias poderia transformar geneticamente bactérias vivas. **(h)** Griffith.

19. Não: diferente de Griffith, que usou cepas IIR e IIIS para demonstrar a transformação, o estudante está usando cepas IIR e IIS. Uma mutação na cepa IIR injetada no camundongo poderia ser suficiente para converter a cepa IIR na cepa IIS virulenta. Ao não fazer o controle adequado de injetar as bactérias IIR apenas, o estudante não pode determinar se a conversão de IIR para IIS se deve a transformação ou a uma mutação. Embora o calor possa ter matado todas as bactérias IIS, o estudante não demonstrou que o calor foi suficiente para matar todas as bactérias IIS. Um segundo experimento de controle útil seria injetar o IIS exterminado por calor em camundongos e observar se qualquer uma das bactérias IIS sobreviveria ao tratamento ao calor.

24. Tubos 1, 4 e 5. O DNA do bacteriófago tem fósforo e a proteína tem enxofre. Quando os bacteriófagos infectam a célula, eles injetam seu DNA na célula, mas os revestimentos de proteína permanecem na superfície da célula. Os revestimentos

de proteína são retirados no liquidificador. Após a centrifugação, os revestimentos de proteína permanecem no líquido, enquanto as células com o DNA estão no fundo do tubo. Assim, as bactérias infectadas com o fago marcado com ^{35}S terão a radioatividade associada com os revestimentos de proteína, enquanto as bactérias com o fago marcado com ^{32}P terão radioatividade associada com as células.

26. O esqueleto de fosfato das moléculas de DNA carreia uma carga negativa, tornando as moléculas de DNA atraentes para o polo positivo da corrente.

27. Aproximadamente 5×10^{20} pares de nucleotídios. Esticado de ponta a ponta, o DNA teria $1{,}7 \times 10^{8}$ km.

29. **(a)**

Organismo e tecido	**(A + G)/(C + T)**	**(A + T)/(C + G)**
Timo de carneiro	1,03	1,36
Fígado de porco	0,99	1,44
Timo humano	1,03	1,52
Medula óssea de rato	1,02	1,36
Eritrócitos de galinha	0,97	1,38
Levedura	1,00	1,80
E. coli	1,04	1,00
Espermatozoide humano	1,00	1,67
Espermatozoide de salmão	1,02	1,43
Espermatozoide de arenque	1,04	1,29

(b) A razão (A + G)/(T + C) de cerca de 1,0 é constante nestes organismos. Cada um deles tem um genoma de fita dupla. A porcentagem das purinas será igual à porcentagem de pirimidinas no DNA de fita dupla, o que significa que (A + G) = (C + T). As proporções (A + T)/(C + G) não são constantes. O número de pares de base A-T relativos ao número de pares de base G-C é único para cada organismo e pode variar entre os diferentes organismos.

(c) A razão (A + G)/(T + C) é quase a mesma para as amostras de espermatozoide, como seria esperado. Como na parte *b*, a porcentagem de purinas será igual à porcentagem de pirimidinas.

32. Adenina = 15%; guanina = 35%; citosina = 35%.

35. O vírus I é um vírus RNA de fita dupla. Uracila está presente, indicando um genoma RNA. Como esperado para um genoma de fita dupla, as porcentagens de adenina e uracila são iguais, assim como as porcentagens de guanina e citosina.

O vírus II é um vírus DNA de fita dupla. A presença de timina indica que o genoma viral é DNA. Como esperado para uma molécula de DNA de fita dupla, as porcentagens de adenina e timina são iguais, assim como as porcentagens de guanina e citosina.

O vírus III é um vírus DNA de fita única. A presença de timina indica um genoma DNA. Entretanto, as porcentagens de timina e adenina não são iguais, assim como as porcentagens de guanina e citosina, o que sugere uma molécula de DNA de fita única.

O vírus IV é um vírus RNA de fita única. A presença de uracila indica um genoma RNA. Entretanto, a porcentagem de adenina não é igual à de uracila e a porcentagem de guanina não é igual à de citosina, o que sugere um genoma de fita única.

36. 100.000.

37. **(a)**

(1) Nenhum dos carbonos 5′ dos dois açúcares está diretamente ligado ao fósforo.

(2) Nenhum dos carbonos 5′ dos dois açúcares tem um grupo OH preso.

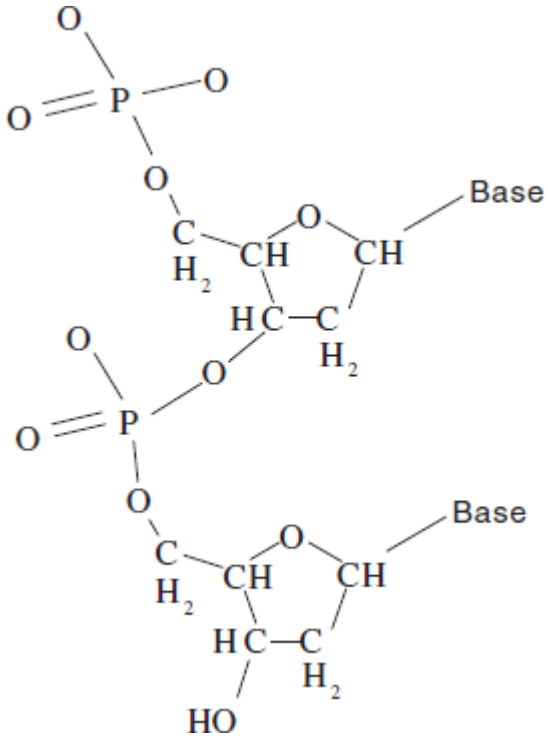
(3) Nenhuma das moléculas de açúcar tem oxigênio na sua estrutura em anel entre os carbonos 1′ e 4′.

(4) Em ambos os açúcares, o carbono 2′ tem um grupo OH preso, o que não ocorre nos desoxirribonucleotídios.

(5) Na posição 3′ em ambos os açúcares, apenas o hidrogênio está preso, oposto a um grupo OH.

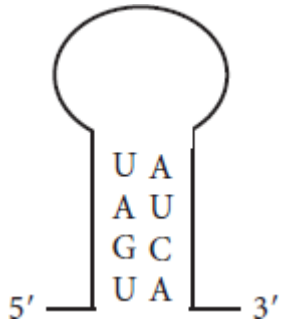
(6) O carbono 1′ de ambos os açúcares tem um grupo OH, oposto a um hidrogênio preso.

(b)



39. Para a estrutura em grampo se formar em uma molécula de RNA, uma sequência de RNA complementar invertida separada por uma região de sequência não complementar é necessária. Os complementos invertidos que formam a estrutura tronco e a alça do grampo é formada pelas sequências não complementares.

    5′-UGCAU-3′ …nucleotídios não pareados… 5′-AUGCA-3′



41. Embora a composição química do material genético no planeta possa ser diferente do material do DNA, é mais provável que ele tenha propriedades semelhantes às do DNA. Como afirmado neste capítulo, o material genético deve ter informação complexa, replicar fidedignamente, codificar o fenótipo e ter a capacidade de variar. Mesmo se o material no planeta não for DNA, ele deve satisfazer estes critérios. Além disso, o material genético deve ser estável.

Capítulo 11

18. Os cromossomos procarióticos em geral são circulares, enquanto os cromossomos eucarióticos são lineares. Os cromossomos procarióticos têm o genoma completo, enquanto cada cromossomo eucariótico tem apenas uma parte do genoma: o genoma eucariótico é dividido em múltiplos cromossomos. Os cromossomos procarióticos são muito menores que os cromossomos eucarióticos e têm apenas uma única origem de replicação do DNA, enquanto os cromossomos eucarióticos têm múltiplas origens de replicação do DNA. Os cromossomos procarióticos estão condensados em nucleoides, que têm alças de DNA compacto em um corpo denso. Os cromossomos eucarióticos têm DNA empacotado em nucleossomos, que são enroscados e empacotados em estruturas de ordem sucessivamente maior. O estado de condensação dos cromossomos eucarióticos varia com o ciclo celular.
19. **(a)** Como cada cromossomo tem duas moléculas de histona H2A e apenas uma molécula de histona H1, as células eucarióticas terão mais H2A do que H1; **(b)** como cada nucleossomo tem duas moléculas de H2A e duas moléculas de H3, as células eucarióticas terão quantidades iguais destas duas histonas.
20. **(a)** Nenhum; **(b)** embrionário; **(c)** adulto; **(d)** nenhum.
22. **(a)** $3{,}2 \times 10^7$; **(b)** $2{,}9 \times 10^8$.

23. Mais acetilação. As regiões de sensibilidade à DNase I são menos condensadas que o DNA que não é sensível à DNase I, o DNA sensível está menos firmemente associado com os nucleossomos, e está em um estado mais aberto. Tal estado está associado com a acetilação dos resíduos de lisina nas caudas de histona do N-terminal. A acetilação elimina a carga positiva do resíduo da lisina e reduz a afinidade da histona pelos fosfatos de carga negativa do esqueleto do DNA.

26. A molécula superior, com uma alta porcentagem de pares de base A-T, terá uma temperatura de fusão menor do que da molécula inferior, que tem principalmente pares de base G-C. Os pares de base A-T têm duas pontes de hidrogênio e assim menos estável que os pares de base G-C, que têm três pontes de hidrogênio.

28. Se a mutação estiver localizada dentro de um gene do cloroplasto, então o gene cloroplasto é herdado da genitora. Esperaríamos os seguintes resultados, não importa qual traço é dominante:

    Macho do trigo selvagem × fêmea do trigo verde-claro → descendentes todos verde-claros.

    Macho do trigo verde-claro × fêmea do trigo do tipo selvagem → descendentes todos do tipo selvagem.

    Se a mutação estiver em um gene nuclear, então ambos os genitores podem transmitir a mutação verde-clara para seus descendentes. Se o tipo selvagem for dominante sobre o verde-claro, então esperaríamos os seguintes resultados:

    Macho do trigo selvagem × fêmea do trigo verde-claro → descendentes todos do tipo selvagem.

    Macho do trigo verde-claro × fêmea do trigo do tipo selvagem → descendentes todos do tipo selvagem.

    Se o fenótipo verde-claro for dominante sobre o tipo selvagem, então esperaríamos os seguintes resultados:

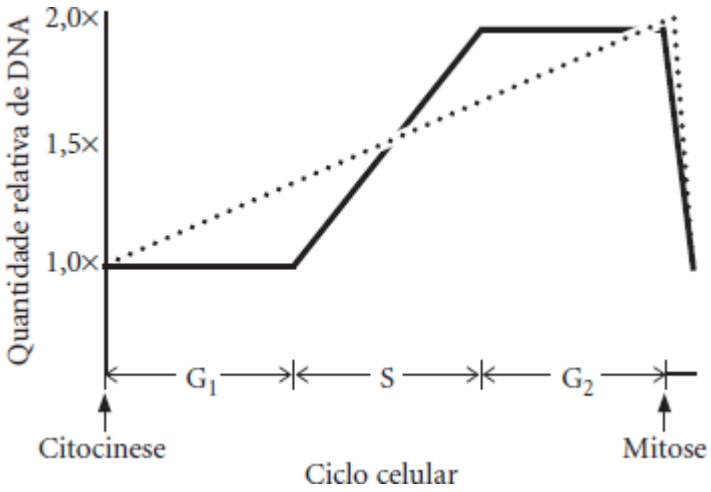
    Macho do trigo selvagem × fêmea do trigo verde-claro → descendentes todos verde-claros.

    Macho do trigo verde-claro × fêmea do trigo do tipo selvagem → descendentes todos verde-claros.

29. O heredograma indica que o distúrbio neurológico é um traço herdado pelo citoplasma. Apenas as fêmeas transmitem o traço para seus descendentes. O traço não parece ser específico para o sexo de modo que os machos como as fêmeas podem ter o distúrbio. Estas características são consistentes com a herança citoplasmática.

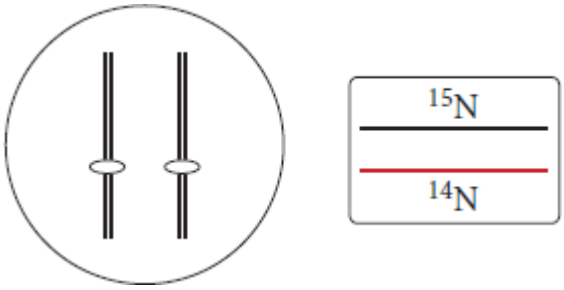
32. Na *Neurospora*, as mutações *poky* são identificadas pelo lento crescimento do organismo mutante. A herança materna indica que as mutações *poky* têm origem mitocondrial porque os genomas mitocondriais são herdados maternalmente na *Neurospora*. Foram identificadas muitas mutações *poky* no genoma mitocondrial. Para uma cepa de *Neurospora* ter uma mutação *poky* com herança biparental, é provável que esta mutação tenha origem nuclear. Na *Neurospora*, os genes nucleares exibem herança biparental. É mais provável que esta mutação *poky* particular afete as vias de produção de energia nas mitocôndrias como indicado pelo fenótipo *poky*, mas a mutação está dentro de um gene nuclear cuja proteína se direciona para as mitocôndrias.

33. Os níveis de DNA nuclear devem aumentar durante apenas a fase S, antes do declínio na citocinese. Os níveis de DNA mitocondrial devem aumentar durante todo ciclo celular, antes de declinar na citocinese.

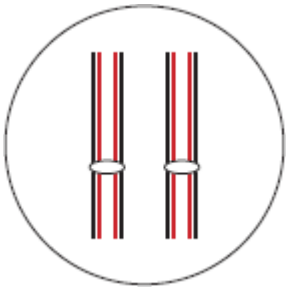


Capítulo 12

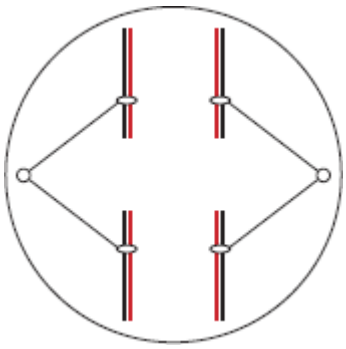
22. **(a)** Células em G_1, antes da troca para o meio com ^{14}N.



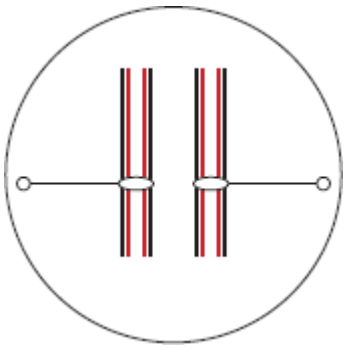
(b) Células em G_2, após a troca para o meio com ^{14}N.



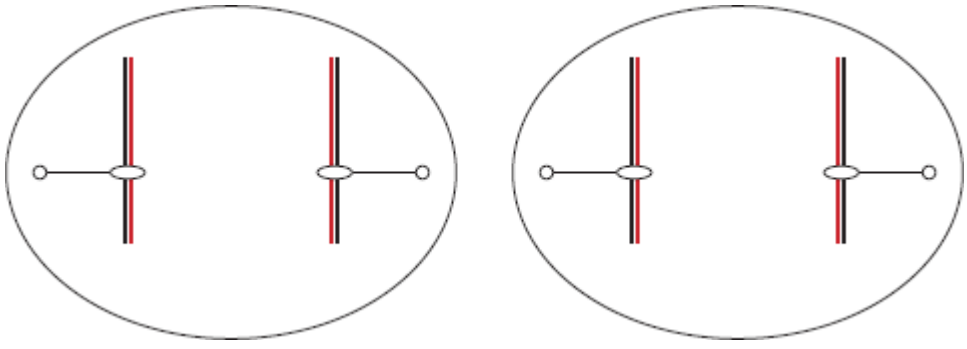
(c) Células na anáfase da mitose, após a troca para o meio com ^{14}N.



(d) Células na metáfase I da meiose, após a troca para o meio com ^{14}N.

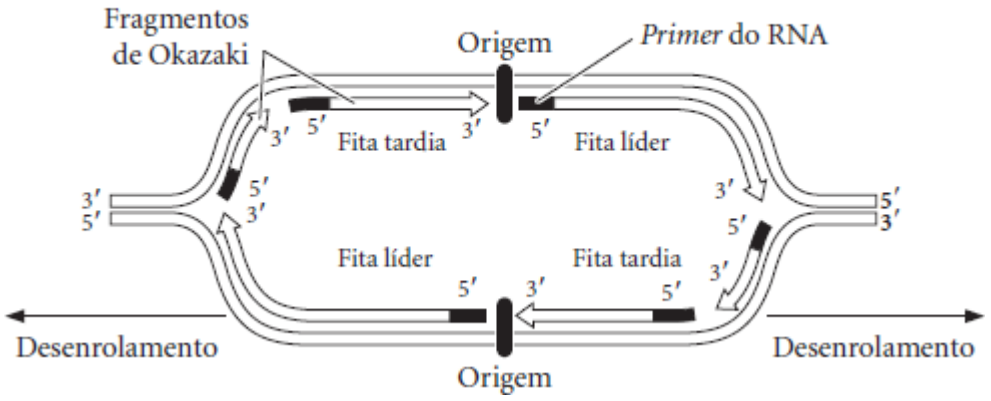


(e) Células na metáfase II da meiose, após a troca para o meio com ^{14}N.



23. Replicação teta, 5 minutos; replicação do círculo rolante, 10 minutos.

25.



27. **(a)** Mais erros na replicação; **(b)** *primers* não seriam removidos;

    (c) *primers* que foram removidos não seriam substituídos.

30. A primase é necessária para o início da replicação na replicação da forma teta. Se a primase não fosse funcional, então não ocorreria o início da replicação, resultando em não replicação.

    A replicação do ciclo rolante não demanda a enzima primase. Uma ruptura de fita única dentro de uma fita fornece um grupo OH 3′ ao qual os nucleotídios podem ser adicionados de modo que a replicação pelo ciclo rolante ocorreria sem uma primase funcional.

33. Devem ser observadas duas bandas distintas dentro do tubo antes da centrifugação. As proteínas histona originais tinham aminoácidos marcados com um isótopo pesado enquanto as histonas recém-sintetizadas têm aminoácidos marcados com um isótopo leve. Se as histonas originais permanecem em uma fita, então esperaríamos observar, após a centrifugação, que as histonas originais sedimentam próximo do fundo do tubo em uma banda distinta. As histonas recém-sintetizadas seriam mais leves e estariam em uma banda distinta superior no tubo.

35. A porção RNA da telomerase é necessária para fornecer o molde para sintetizar as sequências de telômero do DNA complementar nas extremidades dos cromossomos. Uma grande deleção afetaria como os telômeros são sintetizados nas extremidades dos cromossomos pela telomerase e poderiam potencialmente eliminar a síntese do telômero.

40. A proteína B pode ser necessária para o início com sucesso da replicação nas origens de replicação. A proteína B está presente no início da fase S, mas desaparece ao seu fim. A proteína A pode ser responsável por remover ou inativar a proteína B. À medida que os níveis da proteína A aumentam, os níveis da proteína B reduzem evitando eventos de início extras. Quando a proteína A sofre mutação, ela não pode mais inativar a proteína B, assim podem começar ciclos sucessivos de replicação, graças aos altos níveis de proteína B. Quando a proteína B sofre mutação, ela não pode mais auxiliar a iniciação e a replicação para.

Capítulo 13

14. **(a)** É provável que a molécula de RNA seja fita única. Se ela for fita dupla, esperaríamos porcentagem quase iguais de adenina e uracila, assim como porcentagens iguais de guanina e citosina. Nesta molécula de RNA, as porcentagens destes potenciais pares de bases não são iguais, então a molécula é fita única; **(b)** como a fita molde de DNA é complementar à molécula de RNA, esperaríamos porcentagens iguais para as bases na complementaridade de DNA para as bases de RNA. Portanto, esperaríamos no DNA A = 42%, T = 23%, C = 14% e G = 21%.

15. 5′-AUAGGCGAUGCCA-3′

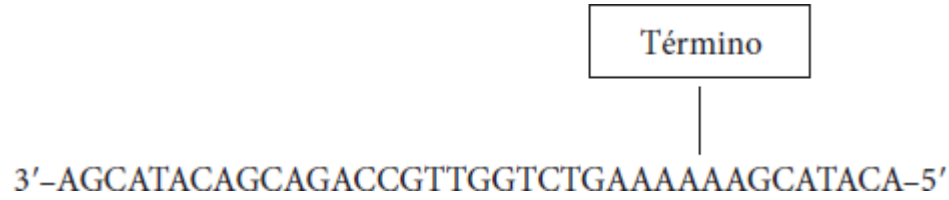
21. A sequência consenso é identificada ao determinar qual nucleotídio é usado com maior frequência em cada posição. Para os dois nucleotídios que ocorrem em uma mesma frequência na primeira posição, ambos estão listados na posição na sequência e identificados por uma barra: T/A G C A A T T.

24.

A região –10 ou Pribnow Box tem a sequência consenso de TATAAT. Entretanto, alguns promotores bacterianos têm a sequência consenso exata.

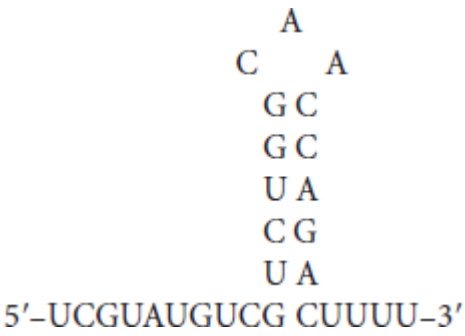
25. **(a)** Afetaria a sequência –10, que seria mais provável de reduzir a transcrição; **(b)** afetaria a ligação do fator sigma ao promotor e regular negativamente à transcrição, reduzindo ou inibindo a transcrição; **(c)** improvável de ter qualquer efeito na transcrição, **(d)** teria pouco efeito na transcrição.

29. **(a)**



(b) Com base na potencial estrutura em grampo que pode se formar e a corrida de U que será sintetizada no RNA, este terminador é um terminador independente de Rho (intrínseco).

(c)



35. Se TBP não pode se ligar a TATA box, então os genes com estes promotores serão transcritos em taxas muito baixas ou não serão transcritos. Como a TATA box é o elemento promotor mais comum para as unidades de transcrição da RNA polimerase II e é encontrada em alguns promotores de RNA polimerase III, a transcrição cai significativamente. É mais provável que a falta de proteínas codificadas por estes genes resulte na morte celular.

39. **(a)** A partir do experimento 3 e sua faixa (*lane*) correspondente no gel, pTFIIB e a RNA polimerase de *S. pombe* (pPolII) parece ser suficiente para determinar o sítio de início da transcrição. O sítio de início da *S. pombe* foi usado mesmo que muitos dos outros fatores de transcrição fossem da *S. cerevisiae*.

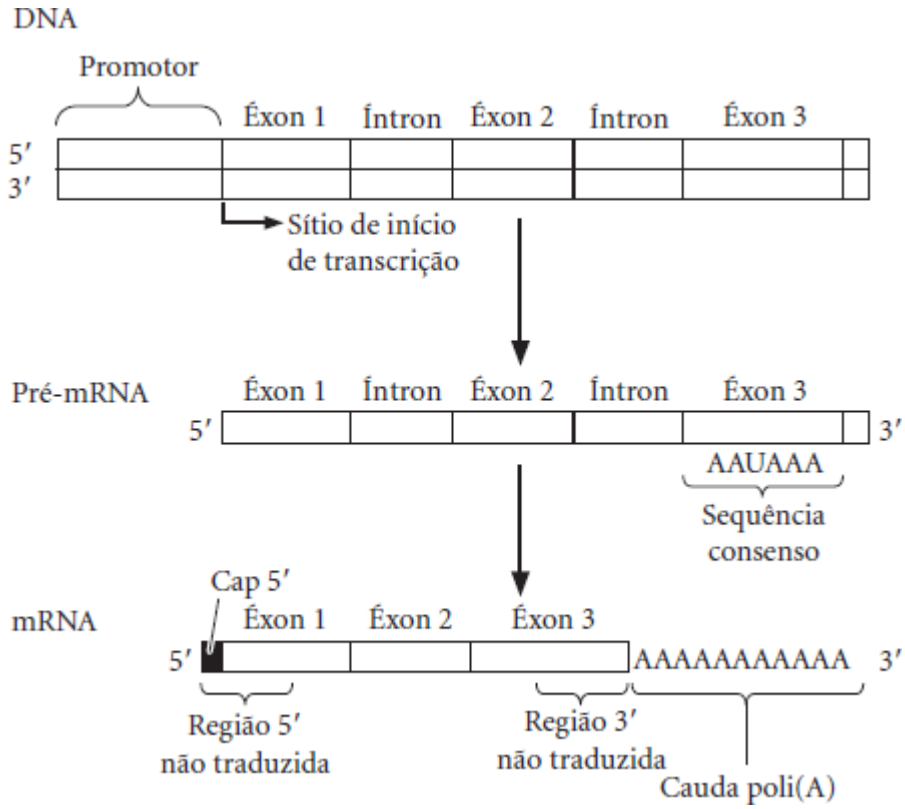
(b) Quando TFIIE e TFIIH foram individualmente trocados, a transcrição foi afetada. Entretanto, a troca pareada de TFIIE e TFIIH possibilitou a transcrição, sugerindo que TFIIE e TFIIH sofrem uma interação espécie-específica e essencial para a transcrição ou que a ausência desta interação inibirá a transcrição. Foram observados resultados semelhantes na troca pareada de TFIIB e RNA polimerase, o que sugere que é necessária a interação espécie-específica entre TFIIB e RNA polimerase para transcrição.

(c) É provável que TFIIB interaja com a RNA polimerase e com outros fatores de transcrição necessários para a iniciação no promotor. Dados indicam que TFIIB e RNA polimerase II têm de ser da mesma espécie para que ocorra a transcrição. A TFIIB potencialmente se associa *downstream* da TATA box, mas em associação com outros fatores de transcrição localizados na TATA box ou próximos a ela. Possivelmente, essas interações *downstream* entre TFIIB e outros fatores de transcrição estimulem mudanças conformacionais na RNA polimerase e na sequência de DNA, permitindo que a transcrição inicie no sítio de início adequado. No experimento 3, RNA polimerase e TFIIB eram de *S. pombe*, levando ao posicionamento da RNA polimerase no sítio de início de transcrição de *S. pombe* pelo pTFIIB.

Capítulo 14

19. É provável que o tamanho grande do gene da distrofina seja por causa da presença de muitas sequências intervenientes, ou íntrons, na região codificadora do gene. A excisão de íntrons por recomposição do RNA gera o mRNA maduro codificador da proteína distrofina.

24.



(a) A região 5′ não traduzida fica *upstream* do sítio de início de tradução. Nas bactérias, o sítio de ligação do ribossomo, ou sequência Shine-Dalgarno, está na região 5′ não traduzida. Entretanto, mRNA eucariótico não tem a sequência equivalente e um ribossomo eucariótico se liga ao cap 5′ da molécula de mRNA.

(b) O promotor é a sequência de DNA reconhecida e ligada pelo aparato de transcrição para iniciar a transcrição.

(c) A sequência consenso AAUAAA, que fica *downstream* da região codificadora do gene, determina a localização do sítio de clivagem 3′ na molécula pré-mRNA.

(d) O sítio de início de transcrição começa na região codificadora do gene e está localizada de 25 a 30 nucleotídios *downstream* da TATA box.

(e) A região 3′ não traduzida é uma sequência de nucleotídios que está localizada na extremidade 3′ do mRNA e não é traduzida em proteínas. Entretanto, não afeta a estabilidade e a tradução do mRNA.

(f) Os íntrons são sequências não codificadoras de DNA que intervêm dentro de regiões codificadoras de um gene.

(g) Éxons são regiões codificadoras de um gene.

(h) Uma cauda poli(A) é adicionada à extremidade 3′ do pré-mRNA; ela afeta a estabilidade do mRNA e a ligação do ribossomo ao mRNA.

(i) O cap 5′ atua na iniciação da tradução e na estabilidade do mRNA.

29. São necessárias proteínas específicas para clivagem da 3′ UTR e para poliadenilação. A mutação dos genes que codificam estas proteínas resultaria em mRNA sem a cauda poli(A) e o mRNA seria degradado mais rapidamente no citoplasma pelas nucleases.

34. **(a)** Na transrecomposição (*trans-splicing*), os éxons de diferentes genes são unidos durante os eventos de processamento de RNA. Essencialmente, o produto mRNA maduro não é produzido pelas sequências de DNA que estão adjacentes ou até necessariamente no mesmo cromossomo. De acordo com o princípio de colinearidade, esperaríamos que a sequência de DNA de um único gene correspondesse à sequência de aminoácidos de uma proteína.

(b) Diferentes mRNAs maduros de um único gene podem ser produzidos por recomposição alternativa. Diferentes proteínas podem ser codificadas dentro do mesmo gene como oposto a um gene correspondendo a uma proteína como

está previsto no conceito de colinearidade.

(c) Na edição do RNA, a informação genética é adicionada ao pré-mRNA depois de ele ser transcrito. O resultado é que a sequência de nucleotídios do gene não corresponde à sequência de aminoácidos da proteína, uma clara violação do conceito de colinearidade.

Capítulo 15

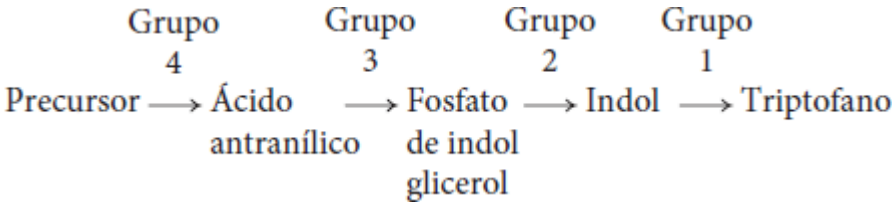
16. Com base na capacidade da cepa mutante em crescer nos substratos fornecidos, as mutações podem ser organizadas em quatro grupos:

    Mutantes do grupo 1 (*trp*-1, *trp*-10, *trp*-11, *trp*-9, *trp*-6 e *trp*-7) podem crescer apenas no meio mínimo suplementado com triptofano.

    Os mutantes do grupo 2 (*trp*-3) podem crescer no meio mínimo enriquecido com triptofano ou indol.

    Os mutantes do grupo 3 (*trp*-2 e *trp*-4) podem crescer no meio mínimo enriquecido com triptofano, indol ou fosfato de indol-glicerol.

    Os mutantes do grupo 4 (*trp*-8) podem crescer no meio mínimo suplementado com a adição de triptofano, indol, fosfato de indol-glicerol ou ácido antranílico.



20. **(a)** 1; **(b)** 2**; (c)** 3; **(d)** 3; **(e)** 4.
21. 3^3 ou 27.
22. **(a)** Amino-fMet-Phe-Lys-Phe-Lys-Phe-carboxila.

    (b) Amino-fMet-Tyr-Ile-Tyr-Ile-carboxila.

    (c) Amino-fMet-Asp-Glu-Arg-Phe-Leu-Ala-carboxila.

    (d) Amino-fMet-Gly-carboxila (o códon de parada UAG acompanha o códon para glicina).
26. **(a)** 3′-CCG-5′ ou 3′-UCG-5′; **(b)** 3′-UUC-5′; **(c)** 3′-AUU-5′ ou 3′-UUU-5′ ou 3′-CUU-5′; **(d)** 3′-ACC-5′ ou 3′-GCC-5′; **(e)** 3′-GUC-5′.
30. Fator de iniciação 3.

    fMet-tRNA^{fMet}.

    Complexo de iniciação 30S.

    Complexo de iniciação 70S.

    Fator de prolongamento Tu.

    Fator de prolongamento G.

    Fator de liberação 1.
33. **(a)** O anticódon 3′ UGC 5′ é complementar ao códon 5′ ACG 3′ que está localizado no sítio A do ribossomo. Observe que o anticódon e o códon são antiparalelos; **(b)** o códon 5′ ACG 3′ codifica o aminoácido treonina.
34. **(a)** A falta de IF-1 reduziria a quantidade de proteína sintetizada. IF-1 promove a dissociação das subunidades maior e menor do ribossomo. A iniciação da transcrição requer uma subunidade menor livre. A ausência de IF-1 reduziria a taxa de iniciação porque mais subunidades menores ribossômicas permaneceriam ligadas às subunidades ribossômicas maiores.

    (b) Não ocorreria tradução. IF-2 é necessário para a iniciação da tradução. A falta de IF-2 evitaria que fMet-tRNA^{fMet} fosse entregue à subunidade ribossômica menor, bloqueando a tradução.

    (c) Embora a tradução fosse iniciada pela entrega de metionina ao complexo ribossomo-mRNA, outros aminoácidos não seriam entregues ao ribossomo. EF-Tu, que se liga a GTP e o tRNA carreado, é necessário para o prolongamento. Este complexo de três partes entra no sítio A do ribossomo. Se EF-Tu não estiver presente, o tRNA carreado não entraria no sítio A, interrompendo a tradução.

(d) EF-G é necessário para a translocação (movimento) do ribossomo ao longo do mRNA na direção 5′ para 3′. Quando é formada uma ligação peptídica entre Met e Pro, a falta de EF-G impediria o movimento do ribossomo ao longo do mRNA e novos códons não seriam lidos. A formação do dipeptídio Met-Pro não requer EF-G.

(e) Os fatores de liberação RF-1 e RF-2 reconhecem os códons de parada e se ligam ao ribossomo no sítio A. Então interagem com RF-3 para promover a clivagem do peptídio do tRNA no sítio P. A ausência de fatores de liberação impediria o término da tradução no códon de parada.

(f) ATP é necessário para os tRNAs sejam carreados com aminoácidos pelas aminoacil-tRNA sintetases. Sem ATP, não ocorreria o carreamento e não haveria aminoácidos disponíveis para a síntese de proteína.

(g) O GTP é necessário para a iniciação, o prolongamento e o término da tradução. Se não houver GTP, não ocorre a síntese de proteína.

38. NMD não deve ser um problema. Acredita-se que NMD seja dependente de proteínas de junção de éxon que normalmente são removidas do mRNA pelo movimento dos ribossomos durante a tradução, mas não são removidos com NMD. Se o primeiro ribossomo a ler o mRNA inserir um aminoácido para o códon de parada devido à ação de PTC124, então ele não deve parar no códon de parada e deve remover qualquer proteínas de junção de éxon. O resultado é que o mRNA será estabilizado, permitindo que ocorra mais tradução.

41. **(a)** O resultado sugere que, para iniciar a tradução, o mRNA é mapeado para encontrar a sequência de início adequada; **(b)** a iniciação da tradução nas bactérias exige 16S RNA da subunidade ribossômica menor interaja com a sequência de Shine-Dalgarno. Esta interação serve para alinhar o ribossomo no códon de parada. Se o anticódon foi trocado de modo que o códon de início não possa ser reconhecido, então é improvável que ocorra a síntese de proteína.

Capítulo 16

11. **(a)** Repressor inativo; **(b)** ativador ativo; **(c)** repressor ativo; **(d)** ativador inativo.

14. A RNA polimerase se liga ao promotor *lac* de forma deficiente, reduzindo muito a transcrição dos genes *lac* estruturais.

19.

	Lactose ausente		Lactose presente	
Genótipo da cepa	**β-Galactosidase**	**Permease**	**β-Galactosidase**	**Permease**
$lacI^+\ lacP^+\ lacO^+\ lacZ^+\ lacY^+$	–	–	+	+
$lacI^-\ lacP^+\ lacO^+\ lacZ^+\ lacY^+$	+	+	+	+
$lacI^+\ lacP^+\ lacO^c\ lacZ^+\ lacY^+$	+	+	+	+
$lacI^-\ lacP^+\ lacO^+\ lacZ^+\ lacY^-$	+	–	+	–
$lacI^-\ lacP^-\ lacO^+\ lacZ^+\ lacY^+$	–	–	–	–
$lacI^+\ lacP^+\ lacO^+\ lacZ^-\ lacY^+/lacI^-\ lacP^+\ lacO^+\ lacZ^+\ lacY^-$	–	–	+	+
$lacI^-\ lacP^+\ lacO^c\ lacZ^+\ lacY^+/lacI^+\ lacP^+\ lacO^+\ lacZ^-\ lacY^-$	+	+	+	+
$lacI^-\ lacP^+\ lacO^+\ lacZ^+\ lacY^-/lacI^+\ lacP^-\ lacO^+\ lacZ^-\ lacY^+$	–	–	+	–
$lacI^+\ lacP^-\ lacO^c\ lacZ^-\ lacY^+/lacI^-\ lacP^+\ lacO^+\ lacZ^+\ lacY^-$	–	–	+	–
$lacI^+\ lacP^+\ lacO^+\ lacZ^+\ lacY^+/lacI^+\ lacP^+\ lacO^+\ lacZ^+\ lacY^+$	–	–	+	+
$lacI^s\ lacP^+\ lacO^+\ lacZ^+\ lacY^-/lacI^+\ lacP^+\ lacO^+\ lacZ^-\ lacY^+$	–	–	–	–
$lacI^s\ lacP^-\ lacO^+\ lacZ^-\ lacY^+/lacI^+\ lacP^+\ lacO^+\ lacZ^+\ lacY^+$	–	–	–	–

21. O gene *lacI* codifica a proteína repressora *lac*, que pode se difundir dentro da célula e se prender a qualquer operador. Ele, portanto, pode afetar a expressão de genes na mesma molécula ou em uma diferente molécula de DNA. O gene *lacO* codifica o operador. Ele afeta a ligação da RNA polimerase ao DNA e afeta a expressão de genes apenas na mesma molécula de DNA.

27. **(a)** Sem expressão gênica; **(b)** transcrição dos genes estruturais apenas quando os níveis de alanina estão baixos; **(c)** sem transcrição; **(d)** sem transcrição; **(e)** a transcrição continua; **(f)** a transcrição continua; **(g)** a transcrição continua.

30. Para bloquear a transcrição, você precisa interromper a ação da RNA polimerase de forma direta ou indireta. O RNA antissenso com sequências complementares ao promotor do gene deve inibir a ligação da RNA polimerase. Se a iniciação da transcrição pela RNA polimerase requer o auxílio de uma proteína ativadora, então o RNA antissenso complementar ao sítio de ligação da proteína ativado do gene também poderia interromper a transcrição. Ao se ligar ao sítio ativador, o RNA antissenso bloquearia o acesso ao sítio pelo ativador e impediria a RNA polimerase de ser auxiliada pelo ativador para iniciar a transcrição.

Capítulo 17

18. É provável que ocorra a floração. A proteína codificada por *FLD* é uma enzima desacetilase. Esta enzima desacetilase remove os grupos acetila das histonas circundantes ao *locus C* da floração (*FLC*). Uma vez que os grupos acetila sejam removidos, a estrutura da cromatina dentro desta região é restaurada. A cromatina restaurada inibe a transcrição a partir do *locus FLC*. O *FLC* codifica um ativador transcricional cuja expressão ativa outros genes que suprimem a floração. Se a transcrição de *FLC* estiver ativa, então não ocorrerá floração.
19. As células do câncer são células de divisão rápida. A metilação do DNA particularmente nas regiões genômicas com muitas sequências CpG (ilhas CpG) estão associadas com a repressão da transcrição). Se as moléculas X31b podem ser captadas pelas células de câncer de rápida divisão, então estas moléculas podem estimular a metilação de sequências de DNA nas células do câncer, levando a repressão transcricional dos genes. A repressão da transcrição poderia retardar o crescimento das células do câncer e provocar a perda da viabilidade estas células.
23. A ação de um acentuador (*enhancer*) é bloqueada quando o isolante está localizado entre o acentuador e o promotor do gene. É provável que os genes A, B e C serão estimulados pelo acentuador e o gene D não será estimulado. Os isolantes bloqueiam a ação estimulante dos acentuadores quando estão localizados entre o acentuador e o promotor do gene. No exemplo da figura, o isolante está apenas entre o gene D e o acentuador. É improvável que o efeito do acentuador nos genes A, B e C seja afetado pelo isolante e estes genes serão estimulados.
24. As moscas-da-fruta desenvolvem: **(a)** características masculinas; **(b)** características masculinas e **(c)** características masculinas e femininas.
26. A presença de elementos ricos em AU está associada com a degradação rápida das moléculas de mRNA que os têm durante um mecanismo de silenciamento de RNA. Se o elemento AU for eliminado, então o miRNA não será capaz de se ligar a sequência consenso do elemento rico em AU e a degradação do RISC não seria iniciada. É provável que esta molécula de mRNA seja mais estável, resultando em maior expressão gênica da proteína codificada pelo mRNA.

Capítulo 18

18. **(a)** Leucina, serina ou fenilalanina; **(b)** isoleucina, tirosina, leucina, valina ou cisteína; **(c)** fenilalanina, prolina, serina ou leucina; **(d)** metionina, fenilalanina, valina, arginina, triptofano, leucina, isoleucina, tirosina, histidina, ou glutamina ou códon de parada também resultaria.
22. **(a)** Uma única substituição de par de bases resultando em uma mutação *missense*.

    (b) Uma substituição de base única resultando em uma mutação *nonsense*.

    (c) A deleção de um único nucleotídio resultando em uma mutação do quadro de leitura.

    (d) Ocorreu uma deleção de seis pares de base, resultando na eliminação de dois aminoácidos (Arg e Leu) da proteína.

    (e) A inserção de três nucleotídios resultando na adição de um códon Leu.
23. Quatro dos seis códons Arg poderiam sofrer mutação por uma substituição de base única para produzir um códon Ser. Entretanto, apenas dois dos códons Arg que sofreram mutação para formar os códons Ser sofreriam mutação em uma segunda posição por uma substituição de base única para regenerar o códon Arg. Em ambos os eventos, as mutações são transversões.

Códon Arg original	Códon Ser	Códon Arg restaurado
CGU	**A**GU	**AGG** ou **AGA**
CGU	**A**GU	**AGG** ou **AGA**

25. Não, a hidroxilamina não pode reverter as mutações *nonsense*. A hidroxilamina modifica os nucleotídios com citosina e pode resultar apenas em mutações de transição GC para AT. Em um códon de parada, a transição GC para AT resultaria apenas em um códon de parada diferente.

27. **(a)** A fita tem duas purinas, adenina e guanina. Como o reparo da depurinação resulta em uma adenina sendo substituída para a purina ausente, apenas a perda de guanina pela depurinação resultará em uma sequência mutante.

    5′-A**G**-3′ para 5′-A**A**-3′

    3′-T**C**-5′ 3′-T**T**-5′

    (b) Pode ocorrer desaminação de guanina, citosina e adenina. Entretanto, é provável que a desaminação de apenas citosina e adenina resulte em sequências mutantes porque os produtos da desaminação podem formar pares de bases inadequados. A desaminação da guanina não pareia com timina, mas ainda pode formar duas pontes de hidrogênio com citosina, então não ocorre mudança.

    5′-**A**G-3′ se A for desaminado, então 5′-**G**G-3′

    3′-**T**C-5′ 3′-**C**C-5′

    5′-A**G**-3′ se C for desaminado, então 5′-A**A**-3′

    3′-T**C**-5′ 3′-T**T**-5′

29. PFI1 provoca transições, PFI2 provoca transversões ou grandes deleções, PFI3 provoca transições e pFI4 provoca inserções ou deleções de base única.

33. A repetição flanqueadora está em negrito.

    5′-ATTCGAAC**TGAC** [elemento de transposição]**TGAC**CGATCA-3′

    5′-ATT**CGAA**[elemento de transposição]**CGAA**CTGACCGATCA-3′

36. Os pares de sequências em (b) e (d) são repetições invertidas porque são ambas invertidas e complementares e podem ser encontradas nas extremidades das sequências de inserção.

38. Estes resultados podem ser explicados pela disgenesia híbrida, com a cepa B abrigando os elementos P e a cepa A sem os elementos P.

42. A aparência de manchas púrpura de tamanhos variados em alguns desses grãos de milho poderia ser explicada pela transposição. Os grãos amarelos podem ser resultado da inativação de um gene de pigmento pela inserção de um elemento *Ds* na planta dessa espiga. Como o elemento *Ds* não pode se transpor por si só, o alelo mutante é estável na ausência de *Ac* e a planta produz grãos amarelos quando fertilizada pelo pólen a partir da mesma cepa (sem *Ac*). Entretanto, alguns grãos podem ter sido fertilizados pelo pólen de uma cepa diferente com um elemento *Ac* ativo. O elemento *Ac* pode mobilizar a transposição do elemento *Ds* do gene do pigmento, restaurando a função do gene do pigmento. A excisão do elemento *Ds* precoce no desenvolvimento do grão produzirá clones maiores de células produtoras de pigmento púrpura. A excisão tardia no desenvolvimento do grão produzirá clones menores das células púrpura.

44. Ao examinar as plantas com níveis aumentados de mutações na sua linhagem celular ou em suas células somáticas. Se elas apresentarem defeito no reparo do DNA, terão maiores taxas de mutação.

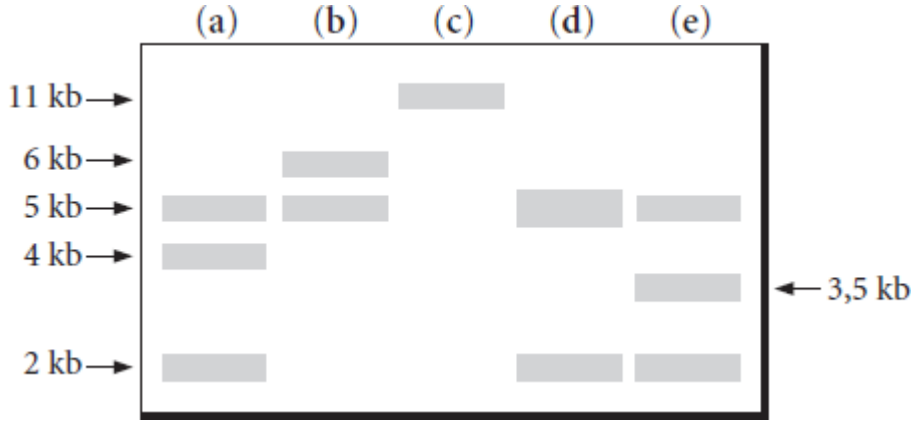
Capítulo 19

25. *Ara*I.

27. 10.

29. **(a)** 460.800; **(b)** 1.036.800; **(c)** 5.120.000.

30.



31. **(a)** Plasmídio; **(b)** fago lambda; **(c)** cosmídio; **(d)** cromossomo bacteriano artificial.

35. Uma biblioteca de cDNA, criada a partir do mRNA isolado de glândula do veneno. As bactérias não podem unir íntrons. Se o engenheiro quer expressar a toxina nas bactérias, então ele precisa de uma sequência de cDNA que foi transcrita de forma reversa a partir do mRNA e, portanto, não tem sequências de íntron. A glândula do veneno deve ser a fonte de mRNA para a síntese de cDNA, de modo que a biblioteca de cDNA será enriquecida para os cDNAs de toxinas.

36. **(a)** Val-Tyr-Lys-Ala-Lys-Trp; **(b)** 128.

39. 5′-NGCATCAGTA-3′

41. **(a)** A mãe de Sally deve ter *A2A3* e *C2C1*. As relações de ligação destes cromossomos são *A1C2* e *A1C3* do pai e *A2C2* e *A3C1* da mãe. A mãe passou um *A2C2* para o irmão de Sally que tem a síndrome G, portanto, o alelo da síndrome G deve estar ligado a *A2C2*. Como Sally herdou o cromossomo *A2C2* de sua mãe, ela também deve ter herdado o alelo da síndrome G, considerando que não ocorre *crossing over* entre os *loci* A, C e G.

    (b) Pai: *A1 C2 g, A1 C3 g*

    Mãe: *A2 C2 G, A3 C1 g*

    Irmão da Sally não afetado: *A1 C3 g, A3 C1 g*

    Irmão da Sally afetado: *A1 C3 g, A2 C2 G*

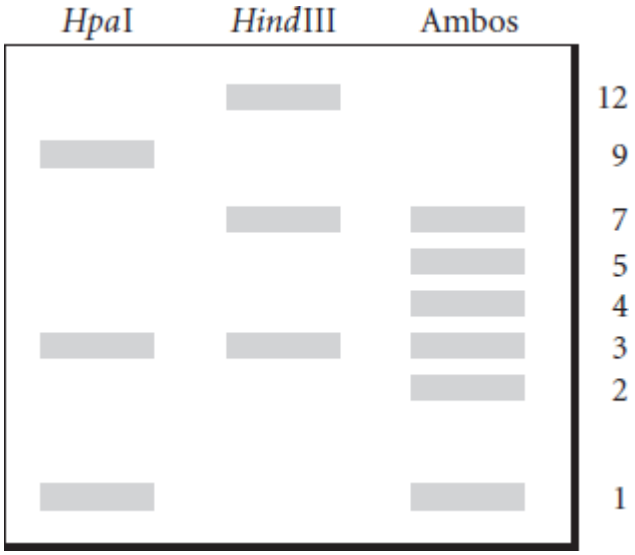
    Sally: *A1 C2 g, A2 C2 G*

43. **(a)** A Figura mostra que ambas larvas e adultos que receberam injeção de dsRNA para *unc22A* expressa altos níveis de proteína *gfp*. Entretanto, as larvas e adultos que receberam injeção com dsRNA para *gfp* não expressa *gfp*. Estes resultados indicam que dsRNA inibe especificamente a expressão do gene correspondente ao dsRNA, mas não a de genes não relacionados.

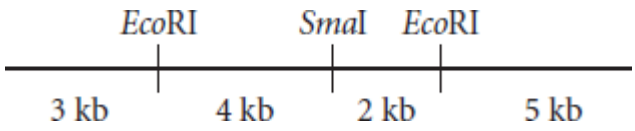
    (b) A injeção de dsRNA correspondente a íntrons e sequências de promotor poderia ter pouco efeito na expressão do gene *gfp* porque a interferência por RNA trabalha ao direcionar o mRNA. Os íntrons e as sequências de promotor não estão presentes no mRNA.

Capítulo 20

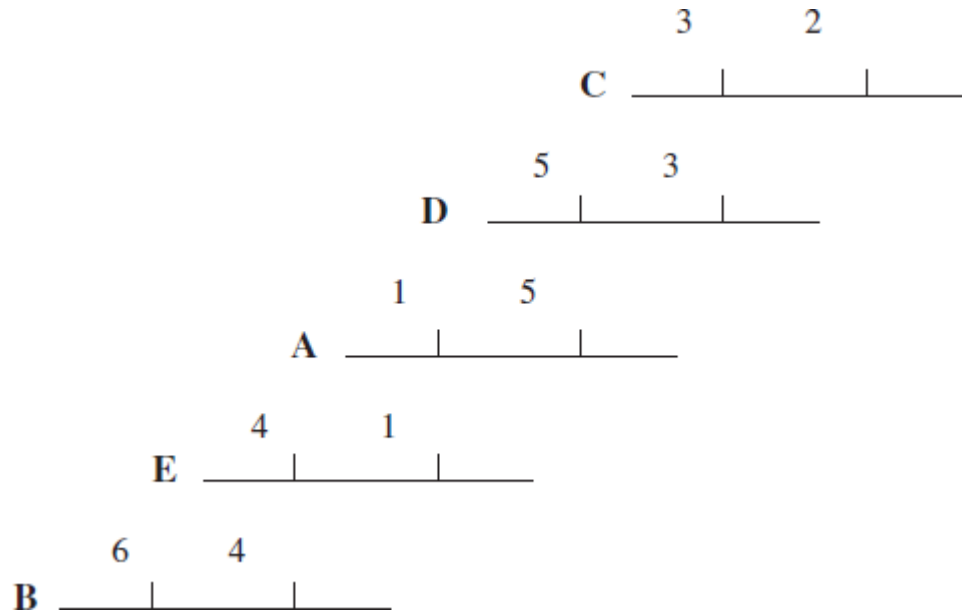
25.



26.



27.



29. **(a)** O cromossomo 22 tem a maior densidade e o maior número de genes, com mais de 500 genes conhecidos e 24 genes novos, enquanto o cromossomo Y tem a menor densidade e menos genes, com menos de 100 genes conhecidos e 23 genes novos.

    (b) Os genes conhecidos encontrados nesta região (0 a 1.000.000 pb) são *PLCXD1*, *GTPBP6*, *PPP2R3B* e *SHOX*. Não foram encontrados genes novos nessa região.

33. Os parálogos, porque ambos se desenvolveram a partir da duplicação de um gene β-globina primordial.

34. Os genes 2 e 24 são expressos em níveis muito maiores nas bactérias resistentes a antibióticos do que as células não resistentes. Por outro lado, os genes 4, 17 e 22 são regulados negativamente. Esses genes podem estar envolvidos na resistência a antibióticos. Os genes regulados positivamente podem estar envolvidos no metabolismo do antibiótico, ou podem executar funções que são inibidas pelo antibiótico. Os genes regulados negativamente podem estar envolvidos na importação do antibiótico ou representam um mecanismo celular que acentua a potência do antibiótico. A caracterização destes genes pode levar a informação sobre o mecanismo de resistência a antibiótico e então ao desenvolvimento de novos antibióticos que podem contornar este mecanismo de resistência.

38. **(a)** 0,5 milhão a mais de 9 milhões de pb nos procariotos; 13 milhões (levedura) a mais de 100 bilhões nos eucariotos; **(b)** Número de genes: De 5.000 a 7.000 nos procariotos; 6.000 a 30.000 nos eucariotos; **(c)** densidade gênica (pb/gene) Aproximadamente 1.000 pb/gene nos procariotos, varia de 2.000 pb/gene para mais de 100.000 pb/gene nos eucariotos; **(d**) número de éxons: com poucas exceções, os genes procarióticos têm zero ou apenas um éxon; os genes eucarióticos multicelulares têm múltiplos éxons.

43. **(a)** O genoma mínimo necessário pode ser determinado ao examinar organismos simples livres tendo pequenos genomas para determinar quais genes eles têm em comum. As mutações podem ser sistematicamente feitas para determinar quais dos genes comuns são essenciais para estes organismos sobreviverem. Os genes aparentemente não essenciais (os genes nos quais as mutações não afetam a viabilidade dos organismos) podem então ser eliminados um por um até ficarem apenas os genes essenciais. A eliminação de qualquer um dos genes essenciais resultará na perda da viabilidade. Por outro lado, os genes essenciais poderiam ser montados por engenharia genética, criando um organismo totalmente novo.

    (b) O organismo novo poderia provar que os humanos adquiriram a capacidade de criar uma nova espécie ou forma de vida. Os humanos então seriam capazes de direcionar a evolução como nunca fizeram. Entre as questões sociais e éticas estaria a pergunta se a sociedade humana tem a sabedoria para moderar seu poder e se tais organismos sintéticos novos podem ou se serão usados para desenvolver patógenos para guerra biológica ou terrorismo. Apesar de tudo, nenhum animal ou pessoa foi exposto previamente ou adquiriu imunidade para tal organismo sintético novo. Também existiria incerteza sobre os efeitos do novo organismo nos ecossistemas se ele for liberado ou escapar.

Capítulo 21

2. Um traço epigenético é um traço estável e transmitido para as células ou descendentes, mas não envolve mudanças na sequência de bases de DNA. Muitos traços epigenéticos são causados por mudanças na expressão gênica resultante de modificações à cromatina. Embora estáveis, muitos traços epigenéticos podem ser influenciados por fatores ambientais.

25. Todas as abelhas-fêmeas desenvolveriam características de rainhas, independente se elas foram alimentadas com geleia real.
26. Não. Nos dinucleotídios CpG, dois nucleotídios citosina em diagonal um do outro em ambas fitas de DNA e ambos estão metilados. A presença de 5-metilcitosina em ambas fitas é necessária para manter a metilação após a replicação. Após a replicação, uma fita de cada nova molécula de DNA está metilada e a outra não. Enzimas metiltransferases especiais reconhecem o grupo metila em uma fita e então metilam a citosina na outra fita, perpetuando o estado metilado do DNA. Nucleotídios citosina individuais não têm um nucleotídio citosina na fita oposta que pode ser metilada após a replicação. Portanto, não são adicionados novos grupos metila pelas metiltransferases após a replicação.
28. Esperaríamos ver diferenças na metilação do DNA e a acetilação da histona que altera a expressão dos genes envolvidos na resposta ao estresse. Também esperaríamos que os adultos apresentassem mais medo e resposta hormonal elevada ao estresse.
29. Como apenas as grávidas foram expostas a vinclozolina, os efeitos sobre o espermatozoide dos camundongos F_2-F_4 não podem ser explicados pelos efeitos diretos da vinclozolina sobre a fertilidade dos machos. Além disso, por causa da alta frequência (90%) dos camundongos afetados em F_2-F_4, parece ser improvável que os efeitos sejam causados pelas mutações induzidas pela vinclozolina. É mais provável que estes efeitos transgeracionais sejam devido a mudanças epigenéticas. Esta ideia é apoiada por diferentes padrões de metilação de DNA dos descendentes F_1-F_4 das fêmeas expostas a vinclozolina. A metilação do DNA é conhecida por afetar a estrutura da cromatina e é responsável por alguns efeitos epigenéticos.
31. **(a)** Ambos os cromossomos X seriam ativos; **(b)** ambos os cromossomos X seriam inativados.

Capítulo 22

14. **(a)** Os produtos do *bicoid* e *dorsal* afetam a polaridade embrionária ao regular a transcrição dos genes-alvo; **(b)** o produto de *nanos* regula a tradução do mRNA *hunchback*.
20. **(a)** As fêmeas com um número maior de cópias do gene *bicoid* teriam níveis maiores de mRNA *bicoid* materno no citoplasma da parte anterior de seus ovos. Após a fertilização, os embriões teriam níveis maiores de proteína Bicoid e estruturas anterior e torácicas aumentadas; **(b)** um número reduzido de cópias do gene *bicoid* resultaria por fim em níveis reduzidos de proteína Bicoid e um embrião com estruturas da cabeça pequenas.
22. A polaridade do ovo (efeito materno): *bicoid, nanos*. *Gap*: *hunchback*, *Krüppel*. Regra de pares: *even-skipped*, *fushi tarazu*. Polaridade de segmento: *gooseberry*.
25. Uma planta sem os genes classe A e B expressaria apenas os genes classe C em todas as quatro pétalas, resultando em flores com apenas carpelos. Uma planta sem ambos os genes das classes B e C expressaria apenas os genes da classe A em todos as quatro pétalas, resultando em flores com apenas sépalas.
27. A expressão difundida de *shh* no primórdio do olho do tetra mexicano provoca a degeneração das células do cristalino. Nestes descendentes F_1 entre o tetra mexicano cego e peixe de superfície, a expressão de *shh* é intermediária entre o tetra mexicano cego e o peixe de superfície, resultando em olhos pequenos.
28. 2.600.

Capítulo 23

17. As deleções podem provocar a perda de um ou mais genes supressores de tumor. As inversões e translocações podem inativar os genes supressores de tumor se os pontos de quebra dos cromossomos estiverem nos genes supressores de genes. Por outro lado, uma translocação pode colocar um proto-oncogene em uma nova localização, onde é ativado por diferentes sequências reguladoras. Finalmente, as inversões e translocações também podem unir partes de diferentes genes, provocando a síntese de uma nova proteína oncogênica.
20. Os retrovírus são fortes promotores. Após sua integração em um genoma hospedeiro, um promotor retrovírus consegue impulsionar a superexpressão de um proto-oncogene celular. Por outro lado, a integração de um retrovírus pode inativar um gene supressor de tumor. Alguns retrovírus carreiam oncogenes que são versões alteradas de proto-oncogenes de hospedeiro. Outros vírus, como o papilomavírus humano (HPV), expressam produtos (proteínas ou moléculas de RNA) que interagem com o maquinário do ciclo celular do hospedeiro e inativam as proteínas supressoras de tumor.
23. Se as diferenças nas taxas de câncer fossem consequentes a diferenças genéticas nas duas populações, então as pessoas que migraram de Utah ou Xangai para outros locais teriam taxas de incidência de câncer semelhantes às das pessoas que

ficaram em Utah ou Xangai. Se as taxas de câncer fossem consequentes a fatores ambientais, então as pessoas que migraram de Utah ou Xangai teriam taxas de câncer determinadas por sua nova localização e não por seu local de origem.

24. **(a)** 50%; **(b)** bilateral; **(c)** o pai pode ter retinoblastoma unilateral, pois a chance aleatória poupou o outro olho da segunda mutação, que teria levado a retinoblastoma nesse olho também.
25. Visto que os oncogenes promovem a proliferação celular, atuam de forma dominante. Em contrapartida, as mutações nos genes supressores de tumor provocam perda da função e atuam de forma recessiva. Quando introduzido nas células, o gene *palladin* mutante aumenta a migração celular. Tal efeito dominante sugere que *palladin* é um oncogene.
32. A região deletada tem um gene supressor de tumor. Os supressores de tumor atuam como inibidores da proliferação celular. A deleção dos genes supressores de tumor possibilitará a proliferação não controlada da célula que é característica do câncer. Os oncogenes, por outro lado, atuam como estimuladores da divisão celular. Assim sendo, a deleção dos oncogenes evita a proliferação celular e, em geral, não provoca câncer.

Capítulo 24

16. **(a)** Característica descontínua, porque existem apenas alguns fenótipos distintos e é determinada por alelos em um único *locus*.

    (b) Característica descontínua porque existem apenas dois fenótipos (anão e normal) e um único *locus* determina a característica.

    (c) Característica quantitativa porque a suscetibilidade é um traço contínuo determinado por múltiplos genes e fatores ambientais (um exemplo de um fenótipo quantitativo com um efeito de limiar).

    (d) Característica quantitativa porque é determinada por muitos *loci* (um exemplo de uma característica merística).

    (e) Característica descontínua porque apenas alguns poucos fenótipos distintos são determinados por alelos em um único *locus*.
17. **(a)** Todos pesam 10 gramas; **(b)** 1/16 pesam 16 gramas, 4/16 pesam 13 gramas, 6/16 pesando 10 gramas, 4/16 pesando 7 gramas e 1/16 pesando 4 gramas.
19. Que seis ou mais *loci* participam.
22. A soma dos pesos é 676, divididos por 10 estudantes, gerando uma média de 67,6 kg.

$$\text{A variância é } s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = 2.024{,}4/9 = 224{,}9$$

$$\text{O desvio padrão} = s = \sqrt{s^2} = 15$$

25. O coeficiente de correlação *r* é calculado a partir da fórmula

$$r = \frac{\text{cov}_{xy}}{s_x s_y}.$$

$$\text{cov}_{xy} = 72$$

$$s_x = 40{,}33$$

$$s_y = 1{,}87$$

$$r = \frac{72}{40{,}33 \times 1{,}87} = 0{,}95$$

27. **(a)** 0,38; **(b)** 0,69.
32. **(a)** 0,75; **(b)** a inexatidão poderia ser devido a uma diferença entre a variância ambiental da população geneticamente idêntica e a da população geneticamente diversa.
35. A única conclusão razoável é **(d)**. A afirmação **(a)** não é justificada porque o valor de herdabilidade não se aplica à altura absoluta nem a um indivíduo, mas à variância da altura nos estudantes da Southwestern University. A afirmação **(b)** não é justificada porque a herdabilidade foi determinada apenas para os estudantes da Southwestern University; estudantes

em outras universidades, com antecedentes étnicos diferentes e de diferentes regiões do país podem ter herdabilidade diferente para altura. A afirmação **(c)** também não se não justifica porque a herdabilidade se refere à variância na altura, e não à altura absoluta. A afirmação **(e)** não é justificada porque a herdabilidade foi determinada para a faixa da variação em fatores não genéticos experimentados pela população em estudo; a variação ambiental fora desta faixa (como desnutrição grave) pode ter efeitos profundos na altura.

36. 0,8.

40. **(a)** Use a equação: $R = h^2 \times S$, em que S é o diferencial da seleção. Neste caso, S = 10 cm – 4 cm = 6 cm e sabemos que a herdabilidade de sentido restrito h^2 é 0,6. Portanto, a resposta à seleção R = 0,6 × 6 cm = 3,6 cm; **(b)** o comprimento médio da asa dos descendentes será o comprimento médio da asa da população mais R: 4 cm + 3,6 cm = 7,6 cm.

45. A largura da cabeça reduzirá. Estes dois traços têm uma correlação genética negativa elevada. Portanto, a seleção para um traço afetará o outro traço de forma inversa.

Capítulo 25

16. $f(T^ET^E)$ = 0,685; $f(T^ET^F)$ = 0,286; $f(T^FT^F)$ = 0,029; $f(T^E)$ = 0,828; $f(T^F)$ = 0,172.

18. Adicionamos todos os alelos X^O ou X^+ e dividimos pelo número total dos alelos X^O e X^+.

    O número de alelos $X^O = 2(X^O X^O) + (X^OX^+) + (X^OY) = 22 + 70 + 36 = 128$

    O número de alelos $X^+ = 2(X^+X^+) + (X^OX^+) + (X^+Y) = 188 + 70 + 112 = 370$

    $f(X^O) = 128/(128 + 370) = 128/498 = 0,26$

    $f(X^+) = 370/498 = 0,74$

22. **(a)** $f(A)$ = (2*42 + 24)/174 = 0,62

    $f(G)$ = (2*21 + 24)/174 = 0,38

    (b) Frequências genotípicas esperadas:

    $f(AA)$ = (0,62)(0,62) = 0,384

    $f(AG)$ = 2(0,62)(0,38) = 0,471

    $f(GG)$ = (0,38)(0,38) = 0,144

    (c)

Genótipo	Observado	Esperado	O-E	$(O-E)^2$	$(O-E)^2/E$
AA	42	33	9	81	2,45
AG	24	41	17	289	7,05
GG	21	13	8	64	4,92

    Qui-quadrado = $\Sigma\ (O–E)^2/E$ = 14,42.

    O número de graus de liberdade é o número de genótipos menos o número de alelos = 3 – 2 = 1.

    O valor p é muito menor que 0,05, portanto, rejeitamos a hipótese de que podem ser esperadas estas frequências genotípicas a partir do equilíbrio de Hardy-Weinberg.

25. Se q = a frequência do alelo de Tay-Sachs, então q^2 = 1/3.600; q = 1/60 = 0,017.

    A frequência do alelo normal = $p = 1 – q$ = 0,983.

    A frequência dos carreadores heterozigotos = $2pq$ = 2(0,983)(0,017) = 0,033, aproximadamente 1 em 30 é carreador.

28. **(a)** 0,64 para M, 0,32 para MN e 0,04 para N; **(b)** com a endogamia, as frequências dos tipos sanguíneos M e N serão 0,648 para M, 0,304 para MN e 0,048 para N.

30. Aqui, usamos a equação $\hat{q} = \frac{\mu}{\mu + \nu}$; em que $\mu = 8 \times 10^{-5}$ e $\nu = 2 \times 10^{-6}$, a frequência no equilíbrio é então $8 \times 10^{-5}/(8 \times 10^{-5} + 2 \times 10^{-6}) = 8/8,2 = 0,98$.

34. **(a)** Como o bandeamento é recessivo, a frequência de cobras listradas = q^2. Na ilha antes da migração, q^2 = 0,4; q = 0,63. No continente, q^2 = 0,81, q = 0,9; **(b)** após a migração,

$$q_{novo} = mq_{imigrantes} + (1 – m)q_{velho} = 0,2(0,9) + 0,8(0,63) = 0,68.$$

35. A pequena população e os efeitos de fundador teriam fortes efeitos da deriva genética. Os geneticistas encontraram uma grande variação nas frequências alélicas entre as populações. Os mesmos fatores acoplados com a endogamia causariam,

nas populações, a perda da variação genética e elevado grau de homozigosidade.

37. **(a)** 0,032; **(b)** 0,145

38. A equação para sobredominância é $q\hat{} = f(AR) = s_{ST/ST}/(s_{ST/ST} + s_{AR/AR})$

    $s_{ST/ST} = 0{,}53$

    $s_{ST/AR} = 0$

    $s_{AR/AR} = 0{,}38$

    $q\hat{} = f(AR) = 0{,}53/(0{,}53 + 0{,}38) = 0{,}53/0{,}91 = 0{,}58$

    No equilíbrio, a frequência de *AR* será 0,58 e a frequência de *ST* será 0,42.

41. **(a)** A aptidão (*fitness*) dos indivíduos com glaucoma é 0,8; o coeficiente de seleção é 0,2.

$$\text{Para um alelo recessivo, } \hat{q} = \sqrt{\frac{\mu}{s}} = \sqrt{\frac{10^{-5}}{0{,}2}} = 0{,}0071.$$

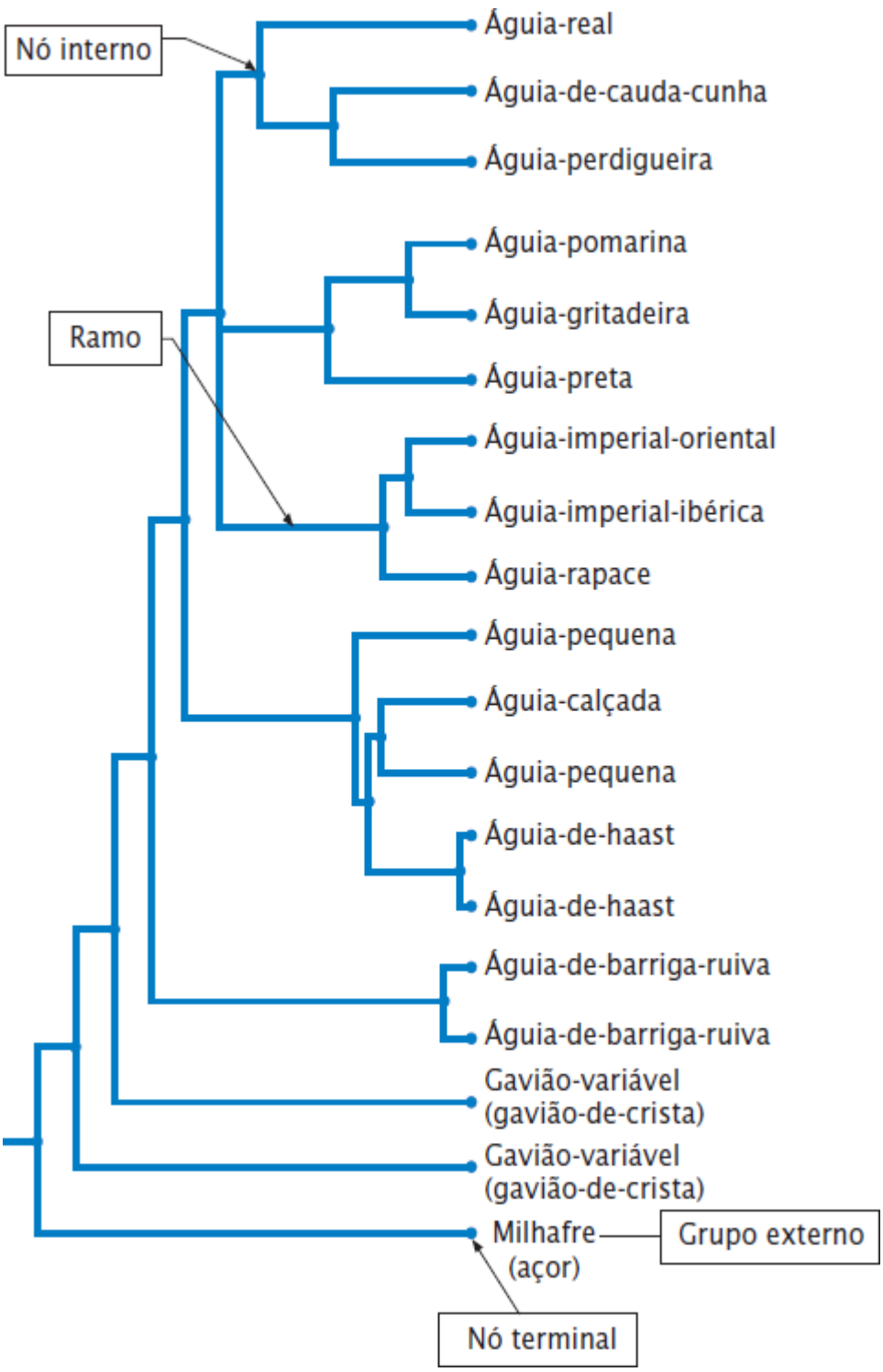
    (b) Frequência de homozigotos = $q^2 = 5 \times 10^{-5}$.

Capítulo 26

21. A primeira ilustração mostra anagênese e cladogênese ocorrendo gradualmente com o tempo. A segunda ilustração mostra que ocorre pouca anagênese (modificação na linhagem) e a maior parte da evolução está associada com a cladogênese, ou seja, a maior parte da evolução ocorre rapidamente quando uma linhagem se divide em duas.

22. Ecológico (diferentes plantas hospedeiras) e temporal (diferentes períodos de acasalamento).

27. **(a)** Os nós terminais são todos grupos taxonômicos listados à direita; **(b)** os nós internos são todos os pontos de ramificação onde as linhagens se dividem; **(c)** os ramos são linhas horizontais que conectam os nós; a linha azul espessa ilustra um exemplo de um ramo; **(d)** o grupo externo é o milhafre (açor), o ramo inferior e o nó na figura.



30. Íntrons. Taxas mais elevadas de substituição são observadas nas regiões gênicas que têm a menor função, porque a seleção natural limita a variação nas partes funcionais dos genes. Embora a região 3′ não traduzida de um gene não codifique os aminoácidos, ela tem sequências importantes na estabilidade e tradução do mRNA. Em um íntron, apenas sequências nas extremidades 5′ e 3′ e no ponto de ramificação atuam na recomposição (*splicing*).